Evaluating the specification errors of asset pricing models ☆

Robert J. Hodrick[a,b,*], Xiaoyan Zhang[a]

[a] *Graduate School of Business, Columbia University, Uris Hall 605b, 3022 Broadway, New York, NY 10027, USA*

[b] *The National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138, USA*

Abstract

This paper evaluates the specification errors of several empirical asset pricing models that have been developed as potential improvements on the CAPM. We use the methodology of Hansen and Jagannathan (J. Finance 51 (1997) 3), and the test assets are the 25 Fama-French (J. Financial Econom. 52 (1997) 557) equity portfolios sorted on size and book-to-market ratio, and the Treasury bill. We allow the parameters of each model's pricing kernel to fluctuate with the business cycle. While we cannot reject correct pricing for Campbell's (J. Political Econom. 104 (1996) 298) model, stability tests indicate that the parameters may not be stable. A robustness test also indicates that none of the models correctly price returns that are scaled by the term premium.

JEL classificaion: C52; G11; G12

Keywords: Hansen-Jagannathan distance; Asset pricing; Time-varying risk prices

☆ We are very grateful to Ken French who provided data for the portfolio returns and the Fama-French factors. We thank Geert Bekaert, Ravi Jagannathan, Martin Lettau, Robert Merton, Sydney Ludvigson, Zhenyu Wang, and seminar participants at Columbia University and the London School of Economics for their comments. We especially thank John Cochrane, the referee, for his insightful comments. Hodrick's research was supported by the National Science Foundation.

*Corresponding author. Graduate School of Business, Columbia University, Uris Hall 605b, 3022 Broadway, New York, NY 10027, USA. Tel.: +1-212-854-3413; fax: +1-212-316-9219.

E-mail address: rh169@columbia.edu (R.J. Hodrick).

Published in: ***Journal of Financial Economics***

1. Introduction

Throughout the 1970s and 1980s, financial economists investigated the pricing implications of the capital asset pricing model (CAPM) developed by Sharpe (1964) and Lintner (1965). The well-known prediction of the CAPM is that the expected excess return on an asset equals the covariance of the return on the asset with the return on the market portfolio times the market price of risk. This price is the ratio of the expected excess return on the market portfolio to the variance of the return on the market portfolio. The expected return prediction of the CAPM can equivalently be stated as the beta of the asset times the expected excess return on the market portfolio, where the beta is the covariance of the asset's return with the return on the market portfolio divided by the variance of the market return.

As empirical research began to uncover a number of expected-return anomalies that the CAPM could not explain, Roll (1977) argued that the model was not testable. Because investors and firms assessing their costs of capital want to know the determinants of expected returns, empirical research continued, but it was necessarily conducted under the recognition that the tests involve a joint hypothesis on the model and the choice of the market portfolio.

The inability of the CAPM to explain the cross-section of asset returns led to the development of a number of alternative empirical asset pricing models. The diversity of these models and the fact that they have been evaluated on a variety of data sets pose severe difficulties for someone who is trying to understand if any of these models is a reasonable replacement for the CAPM. The purpose of this paper is to evaluate and compare a number of these models on a common data set using an appropriate methodology.

Part of our empirical analysis uses the methodology of Hansen and Jagannathan (1997), who develop a distance metric we call the HJ-distance. Hansen and Jagannathan demonstrate how to measure the distance between a true pricing kernel (stochastic discount factor) that prices all assets, and the implied pricing kernel proxy of an asset pricing model. The distance between these two random variables is calculated in the usual way as the square root of the expected value of the squared difference between the two variables. HJ-distance can also be interpreted as the normalized maximum pricing error of the model for portfolios formed from that set of assets. Thus, if the model is correct, the HJ-distance is zero, and there are no pricing errors. Glasserman and Jin (1998) provide an alternative way of comparing models of stochastic discount factors (SDF) by examining the physical probability measures of asset prices and the implied measures of the SDFs. We test whether HJ-distance equals zero using the statistical test developed in Jagannathan and Wang (1996). Because we include a constant in the parameters of the pricing models, we correctly price the average risk-free rate. In this case, the HJ-distance divided by the mean of the pricing kernel is the maximum difference between the Sharpe ratio predicted by the model and the true Sharpe ratio. Consequently, estimation of HJ-distance also provides the maximum expected return error of the model by assuming the investor uses a particular standard deviation.

The models we examine flow from the development of the literature. Even before the CAPM anomalies began to accumulate, theorists such as Merton (1973) noted that the CAPM is a static model, and they developed intertemporal models in which covariances of returns with state variables other than the market return could influence expected returns if the consumption and investment opportunity sets of investors vary over time. Breeden (1979) developed a Consumption CAPM (CCAPM) by demonstrating that an asset's risk premium depends on the covariance of the asset's return with aggregate consumption in continuous time dynamic optimization models. Hansen and Singleton (1982) developed an empirical test of the CCAPM in discrete time by using the Euler equation of the investor's dynamic optimization problem, in which an expected return depends on the covariance of the return with the marginal utility of consumption.

The empirical failure of the CCAPM and the theoretical appeal of the Merton logic led Campbell (1993, 1996) to develop a dynamic asset pricing model in which an expected return depends on the covariances of the return with the market portfolio and with the innovation in the present discounted value of future expected market returns. In the Campbell model, anything that forecasts market returns becomes a risk factor for asset returns.

Jagannathan and Wang (1996) noted that it is possible for the CAPM to hold as a conditional model of expected returns with conditional betas, but the unconditional model would be more complicated since betas could vary over time. They developed an empirical model of this beta-premium sensitivity by taking a stand on the nature of the predictability of market returns.

Cochrane (1996) responded to the failure of the CCAPM by noting that the production side of the economy also must satisfy dynamic Euler equations. This logic led him to develop the implications of a production-based asset pricing model in which covariances of asset returns with macroeconomic measures of investment are important risk factors.

Finally, the empirical failure of the CAPM and the theoretical appeal of multi-factor models led Fama and French (1992, 1993, 1995, 1996) to develop a three-factor model. It is fair to say that this new model, or some extended variant of it, is now the workhorse for risk adjustment in academic circles.

Although the estimation of the parameters associated with the measurement of HJ-distance solves a generalized method of moments (GMM) problem that minimizes a quadratic form based on the average pricing errors from the basic assets, it is not the optimal GMM of Hansen (1982). We also report results from optimal GMM tests of the models, and we generally find similar inference about the validity of the models as in the HJ-distance problems. Neither of these approaches directly minimizes the pricing errors of the basic assets which is equivalent to using an identity matrix in GMM estimation. While such estimation is popular and satisfies the eyes' desire for small errors, inference about the validity of the models is affected severely by the increase in the standard errors associated with this approach. Consequently, we do not report these results.

Because there is considerable evidence that expected returns fluctuate over time, we want to allow for time-varying prices of risks. We do this by allowing the parameters of the models to fluctuate with the business cycle. We measure the business cycle in two ways. One uses the Hodrick and Prescott (1997) filter applied to either industrial production for monthly models or real GNP for quarterly models. The second approach for quarterly models uses the consumption–wealth measure developed by Lettau and Ludvigson (2001a, b). Also, because Loughran (1997) and Daniel and Titman (1997) argue that return characteristics are different in January than outside of January, we use a January dummy variable to allow the parameters of the models to differ across this month and the other months.

Both HJ-distance and optimal GMM assume that the parameters of the model are stable over time. If a model is misspecified because its parameters are not stable, it may nevertheless pass the test of HJ-distance equals zero, but it would not predict well out-of-sample. This situation can characterize both conditional and unconditional models. Ghysels (1998) finds that using conditioning variables to improve asset pricing models may actually worsen their performance out-of-sample because of parameter instability. We therefore follow Ghysels who uses the supLM test developed by Andrews (1993) to investigate instability in parameters.

The common returns that we require each of the models to price are the returns on the 25 portfolios constructed by Fama and French (1993) in which firms are sorted by the market value of their equity (size) and the book-to-market ratio. We use returns in excess of the Treasury bill return, and we also require the models to price the Treasury bill return. The sample period is 1952 to 1997 with either monthly or quarterly data.

Because asset pricing involves conditional expectations, any variable that is in the investors' information set can be used to condition returns. We use this insight to provide a robustness check on the models. The one variable that we use to condition returns is the term spread between the yields on long-term and short-term government bonds.

The paper is organized as follows. The next section provides a discussion of the econometric aspects of the paper including the derivations of HJ-distance, the test that HJ-distance equals zero, and the interpretation of HJ-distance as the maximum difference between the Sharpe ratio of the model and the true Sharpe ratio. Section 3 discusses the data and the parameterization of the different models. Section 4 contains the empirical results. Section 5 provides concluding remarks.

2. HJ-distance and conditional asset pricing models

2.1. Model setup

Assume we have n basic assets to be priced. It is well known that in the absence of arbitrage opportunities there exists a set M of stochastic pricing kernels m which price every asset correctly. That is,

$$\mathrm{E}_t(m_{t+1}R_{j,t+1}) = p_j, \quad \forall j, t > 0, \ \forall m_{t+1} \in M_{t+1}, \tag{1}$$

where m_{t+1} is the stochastic pricing kernel at time $t+1$, M_{t+1} is the set of correct pricing kernels, $R_{j,t+1}$ is the return for portfolio j at time $t+1$, and the price for return $R_{j,t+1}$ at time t is p_j. If $R_{j,t+1}$ is a gross return for a portfolio, then $p_j = 1$; if $R_{j,t+1}$ is an excess return for a portfolio, then $p_j = 0$. The conditional expectation in Eq. (1) is based on the information set at t, denoted Φ_t. By the law of iterated expectations, the unconditional version of Eq. (1) is

$$\mathrm{E}(m_{t+1}R_{j,t+1}) = p_j, \quad \forall j, t > 0, \ \forall m_{t+1} \in M_{t+1}. \tag{2}$$

We use Eq. (2) to estimate and test the various asset-pricing models.

As Hansen and Jagannathan (1997) note, an asset pricing model provides a pricing kernel proxy, y_{t+1}. If the model is true, $y_{t+1} \in M_{t+1}$. We will examine models in which the pricing proxy is a linear function of a constant and a vector of variable factors, f_{t+1}. Define $F'_{t+1} = [1, f'_{t+1}]$, and let the vector of parameters be $b' = [b_0, b'_1]$. Then the pricing proxy is

$$y_{t+1} = b'F_{t+1} = b_0 + b'_1 f_{t+1}, \tag{3}$$

where F_{t+1} is the $k \times 1$ factor vector, and b is the $k \times 1$ coefficient vector. Nonzero elements of b indicate the importance of a factor as a determinant of the pricing kernel. For ease of presentation, we drop the time subscript when it is not necessary for clarity of presentation.

Cochrane (1996) notes that if the model is true, Eq. (2) holds for all n assets with y_{t+1} substituted for m_{t+1}. Then, if p is the $n \times 1$ vector of p_j's, the pricing model has an equivalent representation in terms of multivariate betas and prices of risks:

$$\mathrm{E}(R) = R^0 p + \beta' \Lambda, \tag{4}$$

where $R^0 = 1/E(y)$, $\beta = cov(f, f')^{-1}\, cov(f, R')$, and $\Lambda = -R^0\, cov(f, f')b_1$.

In Eq. (4), R^0 is the unconditional risk-free rate or the zero-beta rate, the β's are the projections of the returns onto the factors, and the Λ's are the prices of beta risks. All of the parameters can be calculated once we know b. To determine whether the jth factor significantly influences the expected returns on a particular set of portfolios, we must assess whether the corresponding Λ_j is significantly different from zero. Notice $\Lambda_j = 0$ does not mean $b_{1,j} = 0$ and vice versa. Only when $cov(f, f')$ is diagonal are the two statements equivalent. The derivations and proofs of these statements can be found in Cochrane (1996).

One must be clear in discussing the prices of factor risks whether it is beta risk or covariance risk. Campbell (1996), for example, uses the covariance decomposition of Eq. (2) to write

$$\mathrm{E}(R) = R^0 p - R^0\, cov(m, R). \tag{5}$$

By substituting the definition of y_{t+1} for m_{t+1} in Eq. (5), one can write

$$\mathrm{E}(R) = R^0 p + \sum_{j=1}^{k} q_j\, cov(f_j, R), \tag{6}$$

where the price of the jth covariance risk is $q_j = -R^0 b_{1,j}$. Since R^0 is not very different from one, we do not report statistics for q_j.

2.2. HJ-distance

Hansen and Jagannathan (1997) note that when the asset pricing model is false, $y \notin M$, and there is a strictly positive distance between y and M. Hansen and Jagannathan define the distance, which we call HJ-distance, as

$$\delta = \min_{m \in L^2} \|y - m\|, \quad \text{where } \mathrm{E}(mR) = p, \tag{7}$$

and the measure of distance is the usual norm, $\|x\| = \sqrt{\mathrm{E}(x^2)}$.[1] The problem defined in Eq. (7) can be rewritten as the following Lagrangian minimization problem:

$$\delta^2 = \min_{m \in L^2} \sup_{\lambda \in R^n} \{\mathrm{E}(y - m)^2 + 2\lambda'[\mathrm{E}(mR) - p]\}. \tag{8}$$

The value of δ is the minimum distance from the pricing proxy y to the set of true pricing kernels M. Let $\tilde{m}$ and $\tilde{\lambda}$ be the solution to Eq. (8). One can think of $y - \tilde{m}$ as the minimal adjustment to y to make it a true pricing kernel. Hansen and Jagannathan (1997) solve Eq. (8) to find

$$y - \tilde{m} = \tilde{\lambda}' R, \tag{9}$$

where

$$\tilde{\lambda} = \mathrm{E}(RR')^{-1}\mathrm{E}(yR - p). \tag{10}$$

[1] Hansen and Jagannathan (1997) also consider a distance measure in which m is required to be strictly positive. If the problem is solved without the constraint and $m_{t+1} > 0$ for all t, the two solutions coincide. In their empirical analysis, Hansen and Jagannathan find this additional restriction does not make a big difference.

Thus, the HJ-distance is

$$\delta = \|y - \tilde{m}\| = \|\tilde{\lambda}' R\| = [\tilde{\lambda}' \mathrm{E}(RR')\tilde{\lambda}]^{1/2}. \tag{11}$$

Substituting for the value of $\tilde{\lambda}$ from Eq. (10) gives

$$\delta = [\mathrm{E}(yR - p)' \mathrm{E}(RR')^{-1} \mathrm{E}(yR - p)]^{1/2}. \tag{12}$$

By solving the conjugate problem to Eq. (8), Hansen and Jagannathan (1997) also provide an important alternative interpretation to δ. It is the maximum pricing error for the set of portfolios based on the basic asset payoffs with the norm of the portfolio return equal to one. We follow Campbell and Cochrane (2000) in interpreting the return errors of the models using this logic.

Consider the return on a portfolio of the n basic assets, $\theta' R$. The true expected return for this portfolio when priced with $\tilde{m}$ is found from Eq. (5) to be

$$\mathrm{E}(\theta' R) = R^0 \theta' p - R^0 \, cov(\tilde{m}, \theta' R). \tag{13}$$

Let $\mathrm{E}^y(\theta' R)$ denote the expected value of the portfolio return predicted by the pricing proxy y. When $\mathrm{E}(y) = \mathrm{E}(\tilde{m}) = (R^0)^{-1}$, we can write

$$\mathrm{E}^y(\theta' R) = R^0 \theta' p - R^0 \, cov(y, \theta' R). \tag{14}$$

By subtracting Eq. (14) from Eq. (13) and using the Cauchy-Schwartz inequality, we have

$$|\mathrm{E}(\theta' R) - \mathrm{E}^y(\theta' R)| = |R^0 \, cov(y - \tilde{m}, \theta' R)| \leqslant R^0 \sigma(y - \tilde{m}) \sigma(\theta' R), \tag{15}$$

where $\sigma(x)$ denotes the standard deviation of x. The inequality in Eq. (15) holds as an equality when the portfolio return is perfectly correlated with $y - \tilde{m}$. Recall from Eq. (9) that $\tilde{\lambda}' R = y - \tilde{m}$, and $\delta = \sigma(y - \tilde{m})$ when $\mathrm{E}(y) = \mathrm{E}(\tilde{m})$. Thus, the portfolio with shares $\theta = \tilde{\lambda}/\delta$ is the maximally mispriced portfolio with norm equal to one. Substituting these results into Eq. (15) and recognizing that $\mathrm{E}(\lambda' R) = 0$ gives

$$\frac{|\mathrm{E}^y(\lambda' R)|}{\sigma(\lambda' R)} = R^0 \delta. \tag{16}$$

The left-hand side of Eq. (16) is the maximum absolute pricing error per unit of standard deviation, or the maximum mispriced Sharpe ratio. Campbell and Cochrane (2000) exploit this idea to evaluate annualized expected return errors of false models by multiplying $R^0 \delta$ by an annualized standard deviation of 20%. We report this type of model return error below.

2.3. Estimation of parameters

Hansen and Jagannathan (1997) note that $\hat{b}$, the estimate of b, can be chosen to minimize δ. To see the relation of this problem to a standard generalized method of moments (GMM) problem, define the pricing error vector $g = \mathrm{E}(yR - p)$, and its sample counterpart

$$g_T(b) = \frac{1}{T}\sum_{t=1}^{T} R_t y_t - p, \tag{17}$$

and let W_T be a sample estimate of $\mathrm{E}(RR')^{-1}$. Then, by squaring Eq. (12), $\hat{b}$ can be chosen as

$$\hat{b} = \arg\min \delta^2 = \arg\min g'_T(b) W_T g_T(b). \tag{18}$$

While Eq. (18) is a standard GMM problem, it is not the optimal GMM of Hansen (1982) which uses as the weighting matrix, $W_T^* = S_T^{-1}$, where S_T is a consistent estimator of $S^* \equiv [T\, var(g_T)]$. Hansen demonstrates that W_T^* is optimal in the sense that the estimated parameters have the smallest asymptotic covariance.

In general, the optimal weighting matrix assigns big weights to assets with small variances in their pricing errors, and it assigns small weights to assets with large variances of their pricing errors. It is obvious that W_T^* changes with different models. This makes it unsuitable for the task of making comparisons among competing models. The alternative weighting matrix of Hansen and Jagannathan (1997) is invariant across competing asset pricing models. Using a common weighting matrix allows us to have a uniform measure of performance across models for a common set of portfolios. The only assumption needed is that the weighting matrix is nonsingular.

Cochrane (1996) argues that $\mathrm{E}(RR')$ may be nearly singular in which case the inversion is problematic, but as we discuss later, we did not encounter inversion problems. To avoid inversion problems and to keep the weighting matrix the same across assets, Cochrane uses the identity matrix as a weighting matrix. This approach is often done in the first-stage estimates of a GMM problem because estimation of W_T^* requires consistent estimates of the parameters.

By assigning equal weights to all basic assets and ignoring cross products of pricing errors, Cochrane's (1996) approach minimizes the sum of squared pricing errors, which is appealing for two reasons. First, it is equivalent to a traditional least squares approach often used in finance, and second, it provides the best graphical representation of predicted returns on the basic assets versus their average returns.

These desirable attributes must be balanced against the theoretical appeal of either optimal GMM or the HJ-distance approach. Optimal GMM provides the most efficient estimates among estimates that use linear combinations of pricing errors as moments. Working with the smallest standard errors provides a more powerful way to test the validity of a particular model. But, because W_T^* is model dependent, it makes no sense to compare chi-square statistics across models. We prefer the HJ-distance approach because it is explicitly designed for comparing the pricing errors of alternative models.

Below we report statistics for both HJ-distance and optimal GMM. We do not report statistics from first-stage estimates because we found them relatively uninformative. Most of the models were not rejected just due to large standard errors, which is economically uninteresting. We also do not find big differences in inference between the results using the two nonidentity weighting matrices.

A big advantages of linear factor models is that they can be solved analytically. To demonstrate the solution, we need to introduce some additional notation. Let the gradient with respect to the parameters be

$$D_T = \frac{\partial g_T}{\partial b} = \frac{1}{T} \sum_{t=1}^{T} R_t F_t'. \tag{19}$$

The analytical solution for $\hat{b}$ from the first order condition of Eq. (18) is given by

$$\hat{b} = (D_T' W_T D_T)^{-1} (D_T' W_T p). \tag{20}$$

From Hansen (1982), the asymptotic variance of $\hat{b}$ is

$$var(\hat{b}) = \frac{1}{T} (D_T' W_T D_T)^{-1} D_T' W_T S_T W_T D_T (D_T' W_T D_T)^{-1}. \tag{21}$$

For optimal GMM, Eq. (21) reduces to

$$var(\hat{b}) = \frac{1}{T} (D_T' S_T^{-1} D_T)^{-1}. \tag{22}$$

One purpose of this paper is to determine whether any of our candidate models of the stochastic discount factor has an HJ-distance equal zero. We construct our test statistics following Theorem 3 in Jagannathan and Wang (1996) as in Appendix A.

We also consider additional model diagnostics. The covariance matrix of the pricing errors for the model is

$$var[g_T(\hat{b})] = \frac{1}{T} \left[I_n - D_T (D_T' W_T D_T)^{-1} D_T' W_T \right] \times S_T \left[I_n - D_T (D_T' W_T D_T)^{-1} D_T' W_T \right]. \tag{23}$$

Thus, we can construct a Wald test statistic for the null hypothesis that $g_T(\hat{b}) = 0$ as

$$g_T'(\hat{b})\, var[g_T(\hat{b})]^{-1} g_T(\hat{b}) \xrightarrow{d} \chi^2(n-k). \tag{24}$$

Since $var[g_T(\hat{b})]$ only has rank $n-k$, we use its pseudo inverse following Cochrane (1996). For optimal GMM, this Wald test reduces to the well-known J-test, with

$$J = g_T'(\hat{b})var[g_T(\hat{b})]^{-1}g_T(\hat{b}) = Tg_T(\hat{b})W_T^*g_T(\hat{b})$$
$$\xrightarrow{\text{d}} \chi^2(n-k). \tag{25}$$

From Eq. (10) the covariance matrix of the Lagrange multipliers is

$$var(\tilde{\lambda}) = W_T\, var[g_T(\hat{b})]W_T. \tag{26}$$

Since the maximum pricing error δ is achieved by $\theta' R$ with $\theta = \tilde{\lambda}/\delta$, we can examine the importance of individual assets to the pricing error by examining the null hypothesis $\tilde{\lambda}_j = 0$.

Finally, it is important to distinguish which pricing errors are under discussion. We defined the pricing errors of the models in Eq. (17). It is the sample average for the differences in prices when we use y to price R minus the correct prices which should be zero for an excess return and one for a gross return. As in other research, we can also define average return errors as

$$\pi = \bar{R} - E^y(R) = \frac{1}{T}\sum_{t=1}^{T} R_t - R^0[p_n - cov(y,R)] = R^0 g_T(\hat{b}). \tag{27}$$

To avoid confusion, we refer to $g_T(\hat{b})$ as model errors and π as the pricing errors of the basic assets. Since R^0 differs slightly across models, the two do not provide the same information. We look at $g_T(\hat{b})$ mainly for details associated directly with δ. We examine π to compare pricing errors for the basic assets across models.

2.4. Conditional models and stability tests

Examining the unconditional implications of linear factor models has two inherent problems. One is that only unconditional risk premiums are estimated. The second is that the models force prices of fundamental risks to be constant across business cycles. Cochrane (1996), Ferson and Harvey (1999), and others try to solve these two problems by using macroeconomic variables as conditioning variables. In Eq. (3), all parameters in b are constant. To allow them to vary with some element z_t in Φ_t, we write

$$\begin{aligned} y_{t+1} &= b'(z_t)F_{t+1} \\ &= (b_{0,1} + b_{0,2}z_t) = [b_{1,1}' + (b_{1,2}z_t)']F_{t+1} \\ &\quad + b_{0,1} + b_{0,2}z_t + b_{1,1}'F_{t+1} + b_{1,2}'(F_{t+1}z_t). \end{aligned} \tag{28}$$

The last equal sign demonstrates Cochrane's point, scaling the prices of factors is equivalent to scaling the factors.

If prices of risks fluctuate over the business cycle, we can capture this effect by using variables that are associated with business cycles. There are three requirements for macroeconomic variables to be legitimate instruments. First, they must be included in the time t information set. Second, they should summarize the status of the business cycle. Third, since the number of the parameters increases geometrically with the number of conditioning variables, which can make the estimates unreliable, the conditioning variables cannot be too numerous. We use only one conditioning variable at a time. Because the previous literature has focused on both monthly and quarterly horizons, we would like a similar conditioning variable for each horizon.

Daniel and Torous (1995) find that the cyclical element in industrial production (IP) is predictive for common stock returns. We adopt their use of IP as one instrument for the monthly models. For quarterly models, we use the cyclical component of real GNP. Because the cyclical components are not observable, we derive both series by using the Hodrick–Prescott (1997) filter applied recursively. We elaborate on the construction of our data in the next section.

Lettau and Ludvigson (2001a) provide an alternative to these output-based measures of the business cycle. Lettau and Ludvigson (2001a) demonstrate that the cyclical element in the log consumption–aggregate wealth ratio (CAY) is strongly predictive for excess stock returns. This argument is consistent with the CCAPM. Lettau and Ludvigson (2001b) test the CCAPM and the CAPM using CAY as a conditioning variable. In their cross-sectional test, conditioning with CAY substantially improves the performance of the models. We also include CAY as a conditioning variable for the quarterly models.

Loughran (1997) and Daniel and Titman (1997) argue that the book-to-market (B/M) effect in stock returns is largely driven by a January effect, that is, the B/M effect is not present at other times of the year. The basic assets we use are the Fama and French 25 portfolios which are constructed precisely to incorporate the B/M and size effects. We use a January dummy variable (JAN) to allow prices of risks to differ between January and other months of the year.

Another important issue is the stability of the model's parameters. Conditional models are attractive because unconditional models may not adequately capture time-varying risk premiums. But, this approach is not costless. If the conditional version is correctly specified and captures the dynamics in risk premiums, it will outperform the unconditional model. However, if the model's implied time-varying risk premiums are inherently misspecified because we choose the wrong conditioning variable, this false model may still appear to work well in small samples since it uses additional degrees of freedom. Ghysels (1998) finds that conditional models are fragile and may have bigger pricing errors than unconditional models.

If the model is correctly specified, parameter stability is not a problem. We use the supLM test of Andrews (1993) to see whether there are structural shifts in the parameters. The null hypothesis is that there are no structural shifts. Andrews argues that the supLM test is powerful against the alternative of a single structural break at an unknown time. He also argues that even if this is not the most interesting alternative hypothesis, it provides a reasonable test of parameter stability. The LM statistics are evaluated at 5% increments between 20% and 80% of the sample, and the largest is the supLM statistic. The distribution for the supLM statistic is presented in Andrews's Table 1.

To keep the estimation tractable, we use the 26 portfolios as the basic assets to be priced. We also investigate whether the model is robust to a different set of assets by adopting Cochrane's approach of scaling returns. Cochrane (1996) notes that conditioning information can be used to scale returns as implied by Eq. (1). These scaled returns can be interpreted as the returns to managed portfolios. The portfolio manager changes the weight of each portfolio according to the signal he observes from the conditioning variable. To illustrate, we multiply both sides of Eq. (1) by any variable $x_t \in \Phi_t$ to get

$$\mathrm{E}_t(m_{t+1} R_{j,t+1})x_t = x_t p_j, \quad \forall j, t > 0, \ \forall x_t \in \Phi_t. \tag{29}$$

By the law of iterated expectations, we have

$$\mathrm{E}(m_{t+1} R_{j,t+1} x_t) = \mathrm{E}(x_t p_j), \quad \forall j, t > 0, \ \forall x_t \in \Phi_t. \tag{30}$$

Eq. (30) provides the orthogonality conditions for scaled returns. If the model is robust to changes in the underlying assets, it should price the new assets correctly. That is, if the model can price nonscaled returns R, under the null hypothesis that the parameters are not asset-sensitive, the model should price scaled returns Rx as well. The test statistic is described in Appendix B.

3. Data

Unless otherwise indicated, all data are from the Center for Research in Security Prices (CRSP). For the monthly models, the sample period in 1952 : 01 to 1997 : 12, for 552 total observations. For the quarterly models, the sample is from 1953 : 01 to 1997 : 04, for 180 total observations. We begin in 1953 : 01 because CAY is only available after 1953 : 01.

3.1. The portfolio returns

Our basic equity assets are the 25 excess returns on the portfolios sorted by size and book-to-market ratio that are calculated as in Fama and French (1993). Excess returns are constructed by subtracting the T-bill rate, and our twenty-sixth asset is the gross return on the T-bill. The previous literature finds that the 25 B/M and size portfolios are very hard to price correctly because they incorporate both size premiums and value premiums. We require the models to price these excess equity returns and the risk-free rate, as well.

Portfolios are numbered 11–55, where the first number refers to the size quintile and the second number refers to the B/M quintile. For example, 11 is the portfolio of the smallest firms with the lowest B/M, while 55 is the portfolio with the largest firms and highest B/M. Table 1 provides summary statistics for the 25 portfolios for the sample period 1952 : 01 to 1997 : 12. It is similar to Table 2 of Fama and French (1993), which involves a shorter sample period from 1963 : 01 to 1991 : 12. For our longer sample, most average returns are larger, except for the low B/M firms. Since the standard errors are smaller, the *t*-statistics are larger except for the low B/M firms. Table 1 indicates that there is considerable difference in the average returns across the 25 portfolios. The average annualized returns range from 4.3% for the smallest firms with lowest B/M ratio to 13.6% for the smallest firms with highest B/M ratio. Within a size quintile, there is a nearly monotonic increase in average returns as B/M increases. Within the B/M quintiles, the average returns to the smallest firms are larger than the average returns to the largest firms, except for the lowest

Table 1
Summary statistics for Fama-French 25 portfolios

The data are monthly returns on the Fama-French 25 portfolios from 1952 : 01 to 1997 : 12 in excess of the one-month T-bill rate. Portfolios are numbered *ij* with *i* indexing size increasing from one to five and *j* indexing book-to-market ratio increasing from one to five.

Portfolios	BM1	BM2	BM3	BM4	BM5
Panel A: Means					
SIZE1	0.36	0.77	0.83	1.03	1.13
SIZE2	0.49	0.78	0.96	1.00	1.15
SIZE3	0.59	0.76	0.80	0.97	1.04
SIZE4	0.60	0.60	0.82	0.87	1.02
SIZE5	0.57	0.63	0.68	0.67	0.85
Panel B: Standard errors					
SIZE1	7.17	6.25	5.56	5.26	5.53
SIZE2	6.49	5.62	5.11	4.85	5.39
SIZE3	5.94	5.04	4.66	4.50	5.14
SIZE4	5.32	4.80	4.61	4.52	5.22
SIZE5	4.54	4.39	4.09	4.24	4.91
Panel C: T-statistics					
SIZE1	1.18	2.91	3.52	4.58	4.82
SIZE2	1.76	3.25	4.41	4.85	5.03
SIZE3	2.33	3.55	4.05	5.04	4.76
SIZE4	2.64	2.93	4.17	4.50	4.60
SIZE5	2.97	3.36	3.89	3.74	4.07

B/M quintile, but there is not monotonicity in average returns across size quintiles.

As demonstrated in Section 2, the weighting matrix for the calculation of HJ-distance depends only on the assets and is the same for different models. The weighting matrix is not the same when we use conditioning information to scale returns. Hence, we have four weighting matrices: monthly and quarterly nonscaled returns, and monthly and quarterly scaled returns. Because our main results are derived from monthly and quarterly nonscaled returns, we focus primarily on these two cases. Eq. (18) demonstrates that the weighting matrix is the estimate of the inverse of the second moment matrix of returns, which must be nonsingular. The condition numbers of the two matrices of sample second moments are 13,548 and 7,851 for monthly and quarterly returns, respectively. For monthly scaled returns, the condition number is 10,264; for quarterly scaled returns, the condition number is 5,238. This indicates that inversion of the matrices should be well behaved.

Cochrane (1996) notes that one can transform the weighting matrix using eigenvalue decomposition such that $W_T = \Gamma Q \Gamma'$ where Γ is an orthonormal matrix with the eigenvectors of W_T on its columns, and Q is a diagonal matrix of eigenvalues. Then, the HJ-distance problem in Eq. (12) can be rewritten as

$$\delta = [\mathrm{E}(yR - p_n)' \Gamma Q \Gamma' \mathrm{E}(yR - p_n)]^{1/2}. \tag{31}$$

The elements of the jth column in Γ can be interpreted as weights that are assigned to the basic assets to form a portfolio associated with the jth eigenvalue in Q. If there are a few large eigenvalues of W_T with eigenvectors that place large weights on only a few portfolios, the GMM problem may be choosing parameters that are associated only with a few portfolios. Because W_T does not change across models, it is fair to ask the competing models to price the same portfolios. But, we do want the structure of the weighting matrix to be reasonable.

Fig. 1 presents the portfolio weights associated with the two largest eigenvalues of the monthly and quarterly weighting matrices. The weights are standardized to sum to one. For monthly returns, Fig. 1 demonstrates that no particular portfolio receives more than twice the weight of the next smallest. Four portfolios, 14, 15, 41, and 42, receive substantial weights, but several other portfolios also receive nontrivial weights. Given that there are other eigenvalues that are also quantitatively important, we conclude that the weighting matrices for the HJ-distance provide a fair challenge to the asset pricing models.

3.2. *Conditioning variables*

We use five variables to capture movements in the prices of risks over the business cycle. For the monthly models, the cyclical part of the natural

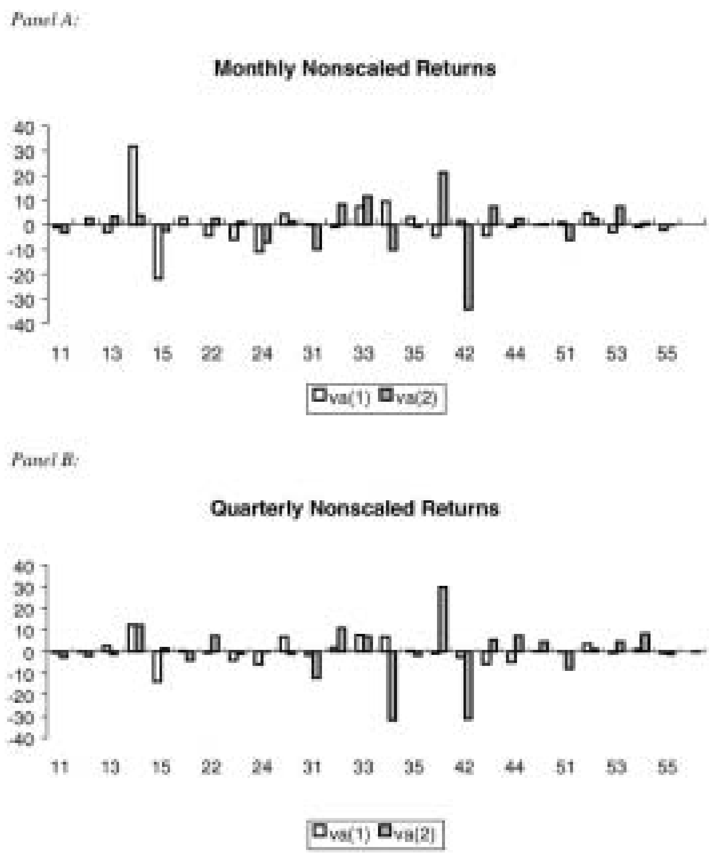


Fig. 1. Standardized eigenvectors of two largest eigenvalues of the weighting matrix $W_T = [(1/T)\sum RR']^{-1}$. The data are monthly and quarterly excess returns of the Fama-French 25 portfolios and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12. Quarterly data are from 1953 : 01 to 1997 : 04. The portfolio numbers on the x-axis are numbered ij with i indexing size increasing from one to five and j indexing book-to-market ratio increasing from one to five. The vector va(1) and va(2) are the eigenvectors corresponding to the two largest eigenvalues of W_T.

logarithm of the industrial production index is one conditioning variable. The industrial production index is from the Citibase monthly data set. The series is available from January 1947 to April 1999. We use the Hodrick–Prescott (1997) filter on the first five years to initialize the cyclical series. The smoothing parameter is set to be 6, 400. Consequently, the first element of our cycle is 1951 : 12. We then use the procedure recursively on all available data to find the subsequent elements for the cyclical series. This method guarantees that each element is in the time t information set. Panel A of Fig. 2 displays the cyclical element of log industrial production index, IP.

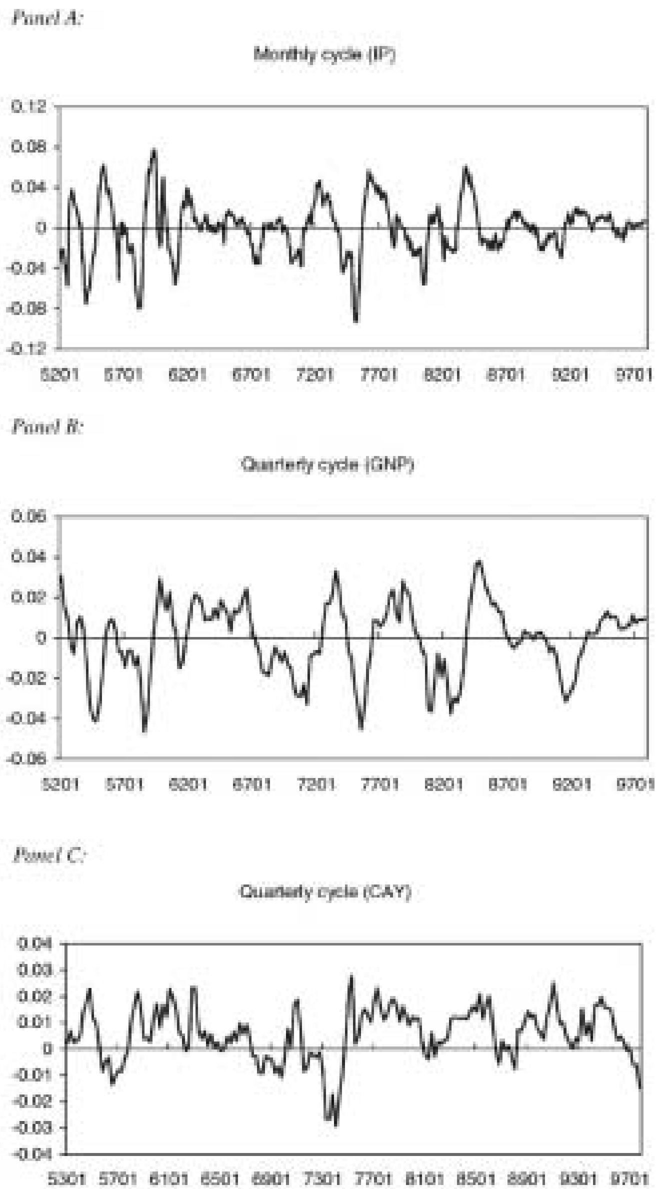


Fig. 2. Time series of three conditioning variables. Cycle (IP) is the cyclical element in monthly Hodrick–Prescott (1997) filtered industrial production. Cycle (GNP) is the cyclical element in quarterly Hodrick–Prescott (1997) filtered GNP. Cycle (CAY) is the aggregate consumption-wealth ratio, derived in Lettau and Ludvigson (2001a). Monthly data for IP are from 1952 : 01 to 1997 : 12. Quarterly data for GNP are from 1952 : 01 to 1997 : 04, and quarterly data for CAY are from 1953 : 01 to 1997:04.

As mentioned above, in monthly models we also scale the factors with a January dummy, JAN, that takes the value one for each January and is zero otherwise. For quarterly models, JAN takes the value one for the first quarter and is zero otherwise.

For the quarterly models, we also scale the factors with the cyclical component of real GNP. The data are also from the Citibase quarterly data set (beginning in 1946 : 01). We use the recursive Hodrick–Prescott (1997) filter with the smoothing parameter equal to 1,600. Because GNP is not announced until the following quarter, we lag GNP once to make sure it is in the time t information set. Alternatively, Lettau and Ludvigson (2001a) develop another conditioning variable, the consumption-wealth ratio, CAY.[2] The CAY series is lagged one period to be a legitimate instrumental variable. Panels B and C of Fig. 2 present the cyclical component of GNP and CAY. The cyclical components of GNP and CAY are not particularly highly correlated. The contemporaneous correlation is -0.0441, and the cross correlations indicate that CAY leads GNP by three to four quarters, as theory predicts consumption should lead income.

Table 2 provides some information on the predictive power of the three conditioning variables except JAN. We use the conditioning variables to estimate the next period return on the value-weighted market return. All of the three conditioning variables have significant predictive power. The explained part of returns is small, as anticipated. With monthly data the R^2 for IP is 1%, and with quarterly data it is 3% for GNP, and 11% for CAY.

We only use one series as the conditioning variables for scaled returns. It is the term premium, calculated as the difference between the 30-year government bond yield and the one-year government bond yield. The data are from CRSP, which provides a monthly index. We construct the quarterly series by using the end-of-quarter observations.

3.3. The asset pricing models

We evaluate eight asset-pricing models. The simplest model incorporates only a constant in the model stochastic discount factor (SDF), and it is called the Null model. The Null model is used as a benchmark. With only a constant factor present, the distance between y and $\tilde{m}$ is $\delta = \min_{m \in M} std(m)$. Thus, we

[2] The data are obtained from Ludvigson's website: http://www.ny.frb.org/rmaghome/economist/ludvigson.html. CAY is calculated as $\text{CAY}_t = c_t - wa_t - (1 - w)y_t$, where c_t is consumption, a_t is asset wealth, y_t is labor income, and w is the weight of asset wealth in total wealth. w is estimated by OLS using all observations. Because of the cointegration relationship between c_t, a_t and y_t, the sample estimate ($\hat{w}$) for w is said to be superconsistent. Lettau and Ludvigson (1999) argue that $\hat{w}$ can therefore be treated as if it is the true parameter. Thus $\widehat{\text{CAY}}_t$, as a function of $\hat{w}$, can be treated as if it is in time t information even though $\hat{w}$ is estimated using all observations, and when using $\widehat{\text{CAY}}_t$ in estimation there is no need to adjust the standard errors for the sampling variability in $\hat{w}$.

Table 2
Predictive power of conditioning variables used to scale factors

The estimated OLS regression is $R_{\text{VW}}(t) = b_0 + b_1$ cycle$(t-1) + \varepsilon(t)$. R_{VW} is the value-weighted return from CRSP. For the monthly regression, the sample period is 1952 : 01 to 1997 : 12. For the quarterly regression, the sample period is 1953 : 01 to 1997 : 04. The series IP and GNP are the Hodrick–Prescott (1997) filtered cyclical components of industrial production and real GNP, respectively. The series CAY is the consumption–wealth ratio calculated by Lettau and Ludvigson (2001a).

	Constant	Cycle	R^2
Panel A: Monthly cycle = IP			
b	0.01	−0.13	0.01
se(*b*)	0.00	0.06	
Panel B: Quarterly cycle = GNP			
b	0.01	−0.77	0.03
se(*b*)	0.01	0.34	
Panel C: Quarterly cycle = CAY			
b	0.00	2.52	0.11
se(*b*)	0.01	0.56	

can interpret the HJ-distance as the standard deviation for the least volatile element in M. In the conditional case, the Null model has two factors, the constant and the conditional *cycle*. The conditional Null model determines whether the movement in the cycle is an important pricing factor.

The second model is the CAPM. The model SDF has two factors, a constant, and the excess return on the market portfolio. We use the return on the value-weighted CRSP index in excess of the one month risk free return, R_{VW}, as a proxy for the excess return on the market. For the quarterly model, we compound the monthly market returns to produce quarterly returns, and we subtract the return on the three-month interest rate. In the conditional model of the SDF, there are four factors: the constant, the *cycle*, R_{VW} and $R_{\text{VW}} \cdot cycle$.

The third model is a linearized CCAPM. The original CCAPM is nonlinear and requires a particular form for the utility function. Rather than develop nonlinear models of marginal utility, we simple use consumption growth, Δc, as the factor. We use the growth rate in real nondurables consumption from Citibase.[3] The unconditional model of the SDF has two factors, the constant and Δc. The conditional model has four factors: the constant, the *cycle*, Δc, and $\Delta c \cdot cycle$.

[3] Martin Lettau and Sidney Ludvigson suggested that the sum of nondurables and services might be a better proxy for consumption. The results with this specification are very similar to our results using only nondurables, except they are a little noisier.

The fourth model is the conditional CAPM developed by Jagannathan and Wang (1996) (hereafter the JW model). The JW model is derived from the assumption that the CAPM holds as a conditional model and that the return on the market is predictable with the default premium, R_{PREM}, which is the difference between the yield on *baa* and *aaa* corporate bonds from the Board of Governors of the Federal Reserve. The JW model's unconditional form involves two betas. One is the original market beta. The other beta incorporates variation in the market beta, which Jagannathan and Wang call beta-premium sensitivity. Beta-premium sensitivity is captured by variation in the default premium. R_{PREM} measures the instability of the market beta over the business cycle. Jagannathan and Wang also argue that the value-weighted index is an inadequate proxy for the market return. They include labor income growth, R_{LBR}, as an additional factor reflecting a return to human capital. Jagannathan and Wang measure labor income growth as $R_{\text{LBR},t} = (L_{t-1} + L_{t-2})/(L_{t-2} + L_{t-3})$, where L is labor income per capita calculated as the difference between personal income and dividend income per capita. The data are obtained from Citibase. Jagannathan and Wang use a two-month average to "minimize the influence of measurement errors." There are consequently four factors in the JW model, a constant, R_{VW}, R_{LBR}, and R_{PREM}. We construct the data as described in Jagannathan and Wang for monthly models. For the quarterly model, R_{LBR} is calculated as the quarterly growth rate in labor income, and R_{PREM} is constructed by selecting the third observation in each quarter. Although the JW model is already an unconditional version of a conditional model, we also estimate our conditional version which implies a total of eight factors in the model SDF.

The fifth model is a linear version of Campbell's (1996) log-linear asset pricing model (described in the tables as CAMP). Campbell (1996) develops an intertemporal asset pricing model that allows for changes in investment opportunities. Factors are determined by their ability to predict the return on the market. As in Jagannathan and Wang (1996), Campbell (1996) argues that labor income is an important additional factor to fully reflect investor's wealth. However, the labor income factor, LBR, is constructed as the monthly growth rate in real labor income (from Citibase). The other three factors are the following: the dividend yield on R_{VW}, DIV; the relative bill rate, RTB, calculated as the difference between the one-month T-bill rate and its one-year backward moving average; and the yield spread between long and short-term government bonds TRM, the difference in yields on the 30-year government bond and on the one-year government bond. In total, there are six factors in the SDF for this model: the constant, R_{VW}, LBR, DIV, RTB, and TRM. In Campbell (1996), the pricing proxy is actually defined as $y = \exp(-F'b)$ and there are constraints across the parameters. Here we simply put the six factors into a linear SDF model, $y = F'b$. For the conditional models, we have twelve factors in total.

The sixth model is a linearized version of Cochrane's (1996) production based asset pricing model (described in the tables as COCH). Cochrane argues that returns should be well priced by the investment return, which is a complicated function of the investment-capital ratio and several parameters. But, Cochrane finds that the investment growth rate performs equally well, and we adopt the investment growth rate model instead of the investment return model. The factors are the growth rate on real nonresidential investment, GNR, and the growth rate on real residential investment, GR. Both original series are from Citibase. The model has three factors in the unconditional model, a constant, GNR, and GR. The conditional Cochrane model has six factors. The data are from Citibase. Since we only have quarterly data for real investment, we do not compute a monthly model in this case.

The above six models are all based on explicit economic theories. We also consider two empirical asset pricing models. They are called empirical because their key pricing factors are derived from the data. The seventh model is the Fama-French (1993) three-factor model (hereafter the FF3 model). The first factor is the excess return on the market portfolio, R_{VW}, as calculated above. To mimic the risk factors in returns related to size and B/M ratio, Fama and French (1993) first sort all stocks into two size portfolios, *small* and *big*, they also sort all stocks into three B/M portfolios, *high*, *medium*, and *low*. Factor SMB (small minus big) is constructed as the difference in returns on *small* and *big*, thus it captures risk related to size. Factor HML (high minus low) is constructed as the difference in returns on *high* and *low*, thus it captures risk related to the B/M ratio. The unconditional model of the SDF has four factors: a constant, R_{VW}, SMB, and HML. We construct quarterly factors by compounding the monthly factors. There are eight factors in the conditional model.

The eighth model is the Fama-French (1993) five-factor model in which they add a termstructure factor and a default-premium factor to their three-factor model (hereafter the FF5 model). The term structure factor, TERM, is the difference between the yield on a thirty-year bond and the yield on the one-month bill. Default risk is the difference between the yields on *baa* and *aaa* corporate bonds (R_{PREM} as in JW). We construct quarterly data by compounding the monthly R_{VW}, SMB and HML, and we use the third observation of each quarter for TERM and R_{PREM}. The conditional model has twelve factors.

4. Empirical results

4.1. Basic model diagnostics

The basic model diagnostics are presented in the seven panels of Table 3. The estimates of HJ-distance are labeled HJ-dist(δ). The p-values of the test $\delta = 0$,

Table 3
Summary of models using nonscaled returns (26 assets)

The data are returns on the Fama-French 25 portfolios in excess of the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12; quarterly data are from 1953 : 01 to 1997 : 04. Cycle (IP) is the cyclical element in the industrial production index; cycle (GNP) is the cyclical element in real GNP; CAY is from Lettau and Ludvigson (2001a). JAN is a dummy variable with value one for January (monthly models) or first quarter (quarterly models) and zero otherwise. HJ-dist(δ) is Hansen-Jagannathan distance. *p*-value for the test $\delta = 0$ calculated under the null $\delta = 0$ is $p(\delta = 0)$. Max. Error is the maximum annual pricing error for a portfolio with annual standard error of 20% under the assumption $E(m) = E(y)$. The standard error for HJ-distance under the alternative hypothesis $\delta \neq 0$ is se(δ). The *p*-value of the optimal GMM test is $p(J)$. The *p*-value of the Wald test that all conditional elements of b^* are zero is *p*-Wald(b^*). The value of the supLM statistics is supLM. An asterisk indicates the model fails the supLM test at the 5% significance level. Number of parameters is No. of para.

MODEL	NULL	CAPM	CCAPM	JW	CAMP	FF3	FF5
Panel A: Monthly models with nonscaled factors							
HJ-dist(δ)	0.420	0.390	0.429	0.386	0.296	0.323	0.316
$p(\delta = 0)$	0.000	0.000	0.000	0.000	0.347	0.000	0.001
Max. Error	8.4%	7.8%	8.6%	7.8%	5.9%	6.5%	6.4%
se(δ)	0.051	0.050	0.063	0.052	0.065	0.052	0.055
$p(J)$	0.000	0.000	0.000	0.000	0.194	0.001	0.005
supLM	216.500*	3.548	4.234	38.290*	193.976*	9.971	58.889*
No. of para	1	2	2	4	6	4	6
Panel B: Monthly models with scaled factors by cycle (IP)							
HJ-dist(δ)	0.410	0.352	0.389	0.314	0.256	0.302	0.273
$p(\delta = 0)$	0.000	0.026	0.041	0.057	0.580	0.010	0.143
Max. Error	8.2%	7.1%	7.8%	6.3%	5.1%	6.1%	5.5%
se(δ)	0.054	0.064	0.084	0.050	0.079	0.062	0.062
$p(J)$	0.000	0.269	0.002	0.062	0.534	0.027	0.218
p-Wald(b^*)	0.006	0.003	0.021	0.016	0.486	0.329	0.398
supLM	10.028	15.963*	9.831	28.254*	73.909*	16.646	40.204*
No. of para	2	4	4	8	12	8	12
Panel C: Monthly models with scaled factors by JAN							
HJ-dist(δ)	0.396	0.366	0.367	0.274	0.284	0.287	0.268
$p(\delta = 0)$	0.000	0.000	0.057	0.650	0.126	0.101	0.335
Max. Error	8.0%	7.3%	7.4%	5.5%	5.7%	5.8%	5.4%
se(δ)	0.060	0.067	0.089	0.086	0.064	0.049	0.067
$p(J)$	0.000	0.000	0.022	0.809	0.065	0.025	0.098
p-Wald(b^*)	0.000	0.165	0.026	0.018	0.962	0.238	0.594
supLM	5.692	6.244	10.345	52.663*	180.979*	13.470	39.225*
No. of para	2	4	4	8	12	8	12

MODEL	NULL	CAPM	CCAPM	JW	CAMP	COCH	FF3	FF5
Panel D: Quarterly models with nonscaled factors								
HJ-dist(δ)	0.649	0.621	0.619	0.578	0.550	0.626	0.537	0.516
$p(\delta = 0)$	0.000	0.000	0.001	0.037	0.016	0.000	0.001	0.018
Max. Error	13.2%	12.6%	12.6%	11.8%	11.2%	12.7%	10.9%	10.5%
se(δ)	0.103	0.097	0.108	0.125	0.107	0.113	0.116	0.105

Table 3 (*continued*)

MODEL	NULL	CAPM	CCAPM	JW	CAMP	COCH	FF3	FF5
$p(J)$	0.001	0.001	0.005	0.083	0.050	0.000	0.010	0.125
supLM	55.023*	3.671	10.071	31.078*	55.957*	10.026	8.746	52.170*
No. of para	1	2	2	4	6	3	4	6
Panel E: Quarterly models with scaled factors by cycle (Lag GNP)								
HJ-dist(δ)	0.642	0.600	0.613	0.543	0.504	0.559	0.452	0.429
$p(\delta = 0)$	0.000	0.001	0.000	0.088	0.147	0.108	0.488	0.362
Max. Error	13.1%	12.2%	12.5%	11.1%	10.3%	11.4%	9.2%	8.7%
se(δ)	0.099	0.082	0.106	0.111	0.104	0.129	0.108	0.099
$p(J)$	0.000	0.011	0.001	0.056	0.101	0.086	0.423	0.254
p-Wald(b^*)	0.219	0.051	0.799	0.013	0.575	0.008	0.111	0.242
supLM	10.837	11.076	11.578	37.006*	44.640*	9.848	11.285	34.071*
No. of para	2	4	4	8	12	6	8	12
Panel F: Quarterly models with scaled factors by CAY								
HJ-dist(δ)	0.634	0.613	0.608	0.544	0.515	0.623	0.528	0.498
$p(\delta = 0)$	0.000	0.000	0.000	0.269	0.099	0.000	0.001	0.011
Max. Error	12.9%	12.5%	12.4%	11.1%	10.5%	12.7%	10.8%	10.1%
se(δ)	0.099	0.110	0.105	0.154	0.125	0.114	0.105	0.090
$p(J)$	0.001	0.000	0.001	0.428	0.097	0.001	0.003	0.032
p-Wald(b^*)	0.012	0.542	0.253	0.404	0.834	0.609	0.931	0.930
supLM	14.028*	14.310	7.170	39.171*	40.373*	16.757	20.149	30.937*
No. of para	2	4	4	8	12	6	8	12
Panel G: Quarterly models with scaled factors by JAN								
HJ-dist(δ)	0.590	0.564	0.582	0.391	0.379	0.510	0.509	0.394
$p(\delta = 0)$	0.001	0.001	0.000	0.997	0.975	0.429	0.005	0.870
Max. Error	12.0%	11.5%	11.9%	8.0%	7.7%	10.4%	10.4%	8.0%
se(δ)	0.135	0.127	0.131	0.239	0.195	0.133	0.129	0.149
$p(J)$	0.011	0.003	0.010	0.997	0.984	0.600	0.004	0.910
p-Wald(b^*)	0.000	0.000	0.006	0.206	0.435	0.001	0.676	0.500
supLM	8.586	9.181	9.133	32.223*	28.311	11.794	20.144	52.123*
No. of para	2	4	4	8	12	6	8	12

as calculated in Appendix A under the null hypothesis that the true distance is zero, are labeled p($\delta = 0$). The maximum annualized expected return error from a portfolio of the basic assets based on Eq. (16) is labeled Max. Error. The maximum pricing error is the product of the HJ-distance and the average risk-free rate times an assumed standard deviation of 20%. The standard errors for the estimates of HJ-distance are labeled se(δ) and are calculated under the alternative hypothesis that the true distance is not equal to zero as in Eq. (45) of Hansen and Jagannathan (1997). These standard errors allow an assessment of the precision with which δ is estimated, and they can thus be used to infer an approximate standard error for the pricing errors in row three by multiplying by the average risk free return and the assumed standard deviation of 20%.

The p-values of the J-statistics from optimal GMM estimates of the models are labeled $p(J)$. The p-values of the Wald tests that the parameters of the scaled factors are all zero are labeled p-Wald(b^*). The values of the supLM tests are labeled supLM, and an asterisk indicates that the test statistic exceeds the 0.05 critical value taken from Table 1 of Andrews (1993). The number of estimated parameters is labeled No. of para.

In finite samples, interpretation of the HJ-distance estimates and their associated maximum pricing errors is hampered by the fact that zero is on the boundary of the parameter space. Even if the null hypothesis is true, in finite samples the estimated HJ-distance will be positive. Of course, if the p-values of the test statistics are well behaved, false rejections of the null hypothesis only occur the correct percentage of the time.

The Monte Carlo experiments conducted by Ahn and Gadarowski (1999) indicate that the expected value of the HJ-distance calculated under the null hypothesis that a three-factor model is true can be quite large and depends on the number of assets and the number of time periods. From Table 1 of Ahn and Gadarowski (1999) with 25 returns, we find average HJ-distances of 0.393 for 160 observations, 0.260 for 330 observations, and 0.174 for 700 observations. Hence, by extrapolating to our monthly sample of 552 observations, we should not be surprised to see an HJ-distance equal to 0.21, even though a three-factor model is true. This corresponds to an annualized maximum pricing error of 4.2%. Similarly, for a quarterly sample of 180 observations, we should not be surprised to see an HJ-distance equal to 0.38 with a maximum pricing error of 7.7%, even though the model is true.

Ahn and Gadarowski (1999) also investigate the empirical size of the test that HJ-distance equals zero. For 25 assets they find that 5.5% of their experiments exceed the 1% critical value with 160 observations, 2.5% are greater with 330 observations, and 1.5% are greater with 700 observations. Thus, for our sample sizes, the monthly model appears to be close to having the correct size of the test if a three-factor model is true, while the rejection rates for the quarterly model appear to be too high.

Panels A–C of Table 3 summarize the results for the monthly models. The first row of Panel A in Table 3 indicates that the Null model, the CAPM, the CCAPM, the JW model, and the FF3 model all have HJ-distances that are larger than or equal to 0.32. The p-values of the tests that these distances are zero are all less than 0.0001. The maximum annualized pricing errors from these models are between 6.5% and 8.6%. The standard errors of the HJ-distances in row four are all about 0.05. Hence, the standard errors of the maximum pricing errors are all about 1%. Generally, we find little disagreement between the Wald tests based on HJ-distance or on optimal GMM of whether the pricing errors on the 26 original portfolios are jointly zero. Consequently, we only report the J-tests from optimal GMM, and in Panel A of Table 3 we find five out of the seven models are rejected at the 0.001

marginal level of significance or smaller. Campbell's model achieves the smallest HJ-distance, and the p-value of the test $\delta = 0$ indicates we cannot reject correct pricing. Thus, the model captures the size and B/M effects and also prices the risk-free rate. It is notable that the same model also passes the J-test. Unfortunately, Campbell's model does not have stable parameters as it fails the supLM test severely.

The HJ-distance of the FF5 model is smaller than that of the FF3 model, but it is still around 0.30. If we subtract the small sample bias in the statistic of 0.21, discussed above, we can conclude that the bias-adjusted HJ-distance is around 0.11 and the maximum annualized pricing error is around 2.2%. As one might suspect, the chief difference between the FF3 model and the FF5 model comes from the fact that the T-bill rate is hard for the FF3 model to price because it only includes equity pricing factors. To evaluate this conjecture, we did a test which only used gross returns on the 25 size and B/M portfolios. There were only small differences between the FF3 model and the FF5 model in that test, and we could reject correct pricing for both models at the 5% marginal level of significance.

Panel B of Table 3 reports the results when the factors of the model SDF's are scaled by cycle(IP). We find the magnitudes of HJ-distances and the corresponding maximum pricing errors all shrink significantly by approximately 10%, except for the Null model. The p-values for the test of HJ-distance equal zero are now between 1% and 5%. We test whether the conditioning information is statistically significant with a Wald test on the joint hypothesis that the parameters for all scaled factors equal zero. For the CAPM, the CCAPM and the JW model, the p-values are smaller than 0.023, which means the scaling variable IP significantly captures time-varying behavior of risks. Using cycle(IP) reduces HJ-distance for all models, and Campbell's model achieves the smallest distance, although there is no significance to the parameters associated with scaling. None of the models pass both the test of HJ-distance equal zero and the supLM test. It is notable that the CAPM with scaled factors marginally passes both the test of HJ-distance equal zero and the optimal GMM test. Again, all results from minimizing HJ-distance are similar to what we find from the optimal GMM approach.

The fact that scaled factor models have smaller HJ-distances than nonscaled factor models comes from two sources. First, the conditioning information reduces the pricing errors by allowing the prices of risks to vary with the business cycle. Second, by doubling the number of parameters, a scaled factor model uses additional degrees of freedom in the minimization problem and is better able to fit the data. This better fit may be spurious, though, as small-sample biases may worsen. The next section examines the details of individual models.

According to Loughran (1997), the January effect explains a substantial part of the B/M effect. When we allow only for a January dummy variable in addition to the constant term of the SDF's, there are very few changes compared to the results in Panel A of Table 3. These results are not reported to save space. Panel C of Table 3 reports results with all factors scaled by JAN. This effectively separates the January observations from the non-January observations by allowing different factor risk prices in January. For the Null model, the Wald statistic for the test that the JAN parameter equals zero is 0.0001, which demonstrates the importance of a January effect. Allowing for a January conditioning variable improves the point estimates of HJ-distance for all the models. Nevertheless, p-values of the J statistics indicated that the CAPM, the CCAPM, and the FF3 models are still rejected at the 0.05 level of significance. The most dramatic improvement is in the JW model which now passes all of the tests except the stability test. The Wald test on the importance of the scaled factors indicates their joint significance. There is a slight improvement in the performance of the FF3 model although the joint test of the significance of the scaled factors has a p-value of 0.15. The FF5 model and Campbell's model already do reasonably well with nonscaled factors. Scaling all the factors in these models with a January dummy does not appear to add any important factors since the p-values of the Wald tests are both quite large.

The previous literature typically reports either monthly or quarterly models. Some models, such as Cochrane's (1996) model, can only be applied to quarterly data because of data constraints. In this section we investigate the performance of the models with quarterly data. Several issues arise. First, time aggregation may worsen the fit between the factors and the models by smoothing the factors.[4] Second, market imperfections that cause short-term deviations from the models may be lessened because the returns are cumulated. Third, as noted above, the small-sample performance of any model deteriorates with a smaller number of observations. The first and third effects suggest the performance of the models with quarterly data deteriorates, while the second factor allows for improvement.

Panel D provides the summary results for the eight quarterly models, the seven previously investigated plus Cochrane's (1996) model. Although the point estimates of the HJ-distances are much larger for the quarterly models than the monthly models, recall from our discussion of Ahn and Gadarowski (1999) that values like 0.38 are to be expected in these sample sizes even if a three-factor model is true. Nevertheless, the quarterly HJ-distances generally exceed the average of the Ahn and Gadarowski figures by more than the monthly estimates exceed the corresponding average from the Monte Carlo experiments. For example, the monthly FF3 models has an HJ-distance of 0.323 and the Monte Carlo average is approximately 0.21 for a difference of

[4]This logic leads Cochrane (1996) to time average monthly returns in constructing quarterly returns. While we construct the quarterly returns from the compound monthly returns as $R_{t+1} + R_{t+2} + R_{t+3}$, Cochrane (1996) uses $\frac{1}{3}R_{t+1} + \frac{2}{3}R_{t+2} + R_{t+3} + \frac{2}{3}R_{t+4} + \frac{1}{3}R_{t+5}$.

0.113. At the quarterly sampling interval we find a difference of $0.537 - 0.38 = 0.157$. Using this bias-adjusted value to calculate the maximum pricing error for the FF3 model leads to a value of 3.2% rather than the 10.9% reported in Panel D.

While the *p*-values of the tests that HJ-distance equals zero are all less that 0.037, recall also that in this sample size the asymptotic *p*-values probably understate the probability of a Type I error as Ahn and Gadarowski (1999) find that 15.7% of their empirical experiments exceed the 5% asymptotic critical value in samples of 160 observations. Hence, it seems reasonable to conclude that the evidence against the JW model, the FF5 model, and Campbell's model is not particularly strong. Unfortunately these three models all fail the parameter stability test.

In Panel E, we scale all factors by the lagged cyclical component of GNP. Including this conditioning information reduces the magnitudes of HJ-distance and the associated maximum pricing errors by 5–10%. Two models, the FF3 model and Cochrane's, now pass the test of HJ-distance equal zero and the supLM test, although Cochrane's model has a considerably larger δ. Once again the HJ-distance tests are consistent with the results from optimal GMM. The tests that all parameters for scaled factors equal zero indicate scaling with GNP does not significantly improve the performance of the models. One should keep in mind, though, this is a joint test which may overshadow the significance individual parameters.

An alternative quarterly scaling variable is the consumption-wealth ratio, CAY, from Lettau and Ludvigson (2001a). They find that scaling with CAY greatly improves the performance of the CCAPM in pricing the excess returns on the 25 Fama-French portfolios over a sample period 1963–1997 when the returns are equally weighted. However, evaluating the model with the HJ-distance metric for our sample of 1953 to 1998 indicates that scaling with CAY does not produce a noticeable improvement for the CCAPM. The scaled model fails both the test of HJ-distance equal zero and the optimal GMM test. None of the models scaled by CAY passes both the test of HJ-distance equal zero and the supLM test. The Wald test of the importance of the scaling parameters also does not indicate strong statistical significance of CAY.

Panel G provides results when all the factors are scaled by JAN. For the quarterly models, JAN takes the value one for the first quarter of each year and the value zero otherwise. The first thing to note is scaling all factors with JAN reduces the magnitude of the HJ-distance for all models. The JW model, Campbell's model, and the FF5 model all have *p*-values for the test of HJ-distance equal zero above 80%. The annualized pricing errors for these three models also are now less than or equal to 8%, which is in the range of correct pricing given the bias discussed above. Surprisingly, the FF3 model does not pass the HJ-distance test and the J test. This is because the scaled factor model is still unable to price the small growth firms. Cochrane's model passes both the test of HJ-distance equal zero and the supLM test. More details for this model are provided in the section on successful models.

4.2. Model errors for nonscaled factor models

Additional information on the performance of the models is available by examining the model errors and the Lagrange multipliers which are the components of δ. To check whether conditioning information improves the performance of a model, we first need to understand the performance of the original nonscaled factor model. The average model errors from HJ-distance estimates with a two standard error band are presented in Fig. 3. Since monthly unconditional model errors share very similar patterns with the quarterly model errors, we only present monthly model errors (g_T) as defined in Eq. (17). For Cochrane's model, we report quarterly model errors.[5]

In Panel A of Fig. 3, the model errors for the Null model range from -0.01% for the T-bill to 1.15% per month for portfolio 25. Remember that the first number of a portfolio indexes the size quintile with increasing numbers indicating increases in size and that the second number of a portfolio indexes the book-to-market ratio with increasing numbers indicating increases in B/M. The B/M effect is very evident in Fig. 3 as in each size quintile, higher B/M portfolios have larger average pricing errors. As we increase across size quintiles, there is less dispersion in the pricing errors but no particularly pronounced decrease in average pricing errors. The model under-estimates the returns on all portfolios except the T-bill rate.

Panel B of Fig. 3 demonstrates that the CAPM correctly prices the largest size portfolios, but it tends to under-estimate returns on high B/M portfolios and to over-estimate returns on low B/M portfolios. The model errors range from -0.50% per month for portfolio 11 to 0.45% per month for portfolio 15.

The CCAPM is presented in Panel C of Fig. 3. It has a pattern very similar to the Null model, which is consistent with the correlation of 0.93 between the adjustment, $y - \tilde{m} = \tilde{\lambda}' R$, to the Null model and the adjustment to the CCAPM to make it a correct SDF. High B/M firms are more severely underpriced by the CCAPM than by the CAPM.

The JW model is presented in Panel D of Fig. 3. It has a very similar pattern to the CAPM except the over-estimation for low B/M portfolios is slightly smaller. This is not surprising in light of the correlation of 0.99 between the adjustments to the CAPM and to the JW model.

Panel E of Fig. 3 reports the pattern for Campbell's pricing errors. The model considerably attenuates the B/M effect. The average errors range from

[5]We also examined model errors from minimizing the equal-weighted sum of squared pricing errors, that is using an identity matrix as the weigthing matrix. The patterns of errors across the various models are quite similar to the errors in Fig. 3 and are consequently not reported.

−0.28% to 0.30%. Part of the ability of the model to pass the test of HJ-distance equal zero arises from its increased standard errors relative to the CAPM. Although δ can be compared across models, the p-values of the tests are not comparable because they are based on the eigenvalues of A in Appendix A which depend on the pricing factors, the variance of pricing errors, and the number of parameters.

Panel F presents the pricing errors in Cochrane's quarterly model which shares the same magnitude and pattern as the quarterly CAPM, which is not presented. There is a distinct B/M effect as in the monthly CAPM. The correlation between the adjustment to Cochrane's model to make it a correct pricing model and the adjustment to the quarterly CAPM is 0.97.

The FF3 model is presented in Panel G. The presence of the two factors SMB and HML in addition to the market return considerably dampens the B/M effect present in Panel B. Now there is no particular pattern for the model errors. They are scattered around the zero axis. The FF3 model overpredicts the average returns for both the smallest firms and the largest firms, but especially the small growth stocks (smallest firms with low B/M ratios). The FF5 model in Panel H has a similar pattern to the FF3 model, except it reduces the pricing errors slightly. The correlation of the adjustments to the two models is 0.98.

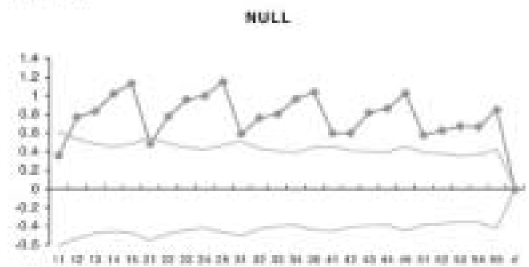
All models share one common characteristic, they do not misprice the T-bill rate. Model errors for the T-bill rate are always around zero.

4.3. Interesting models

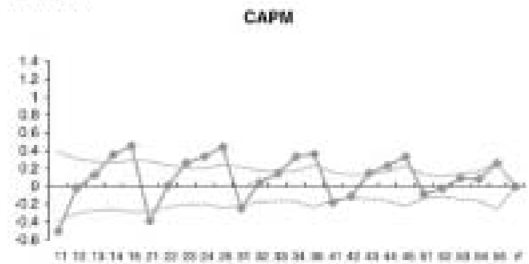
Since we have 21 monthly models and 32 quarterly models, we cannot display all the parameter estimates, but we report results for "interesting models". We define "interesting" as a model that at least marginally passes the test of HJ-distance equal zero at the 1% marginal level of significance. We also require that the scaling parameters for an interesting scaled factor model are jointly significant at the 5% level. Because inference about the validity of the models based on the test of HJ-distance equal zero is always similar to inference based on the J test from optimal GMM, passing the J test is implicitly also a criterion. In total we have 12 models satisfying both

Fig. 3. Model errors for monthly models with nonscaled factors. The data are monthly and quarterly excess returns of the Fama-French 25 portfolios over the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12. Quarterly data are from 1953 : 01 to 1997:04. The portfolio numbers on the x-axis are numbered ij with i indexing size increasing from one to five and j indexing book-to-market ratio increasing from one to five. The diamonds are the model errors, as defined in Eq. (17), and the numbers are in monthly (quarterly from Cochranes's model) percent. The two other lines provide a two standard error band.

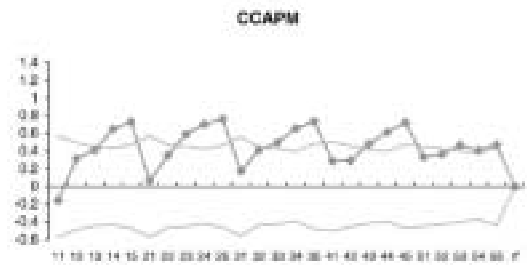
Panel A:



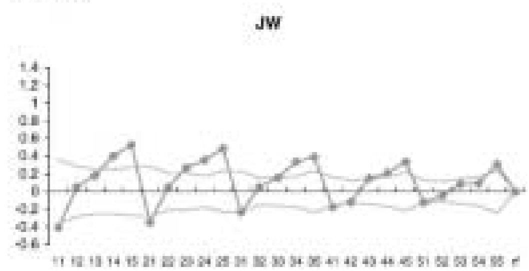
Panel B:



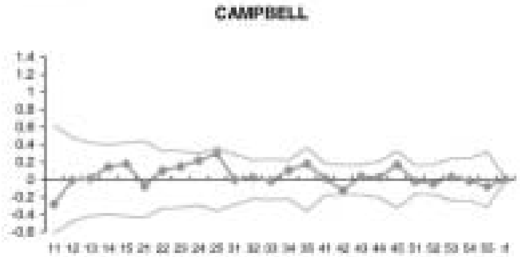
Panel C:



Panel D:

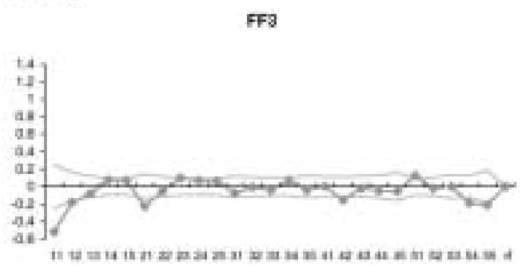


Panel E:



Panel F:

Panel G:



Panel H:

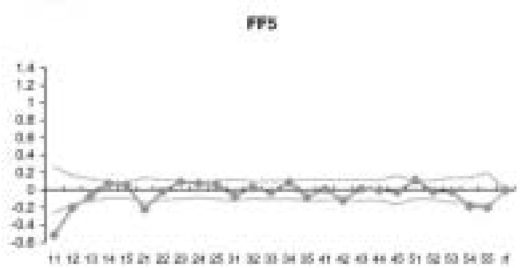


Fig. 3. (*continued*)

conditions. In addition we provide information on the monthly FF3 model with nonscaled factors for comparison. This section first discusses monthly models, then quarterly models.

Table 4 reports parameter estimates from minimizing the HJ-distance measure for the interesting models. Each panel has two parts. The first part presents estimates for b as in Eq. (3). If b_1 for one factor is significantly different from zero, then that factor is an important determinant of the pricing kernel. The second part of each panel presents estimates for the prices of risks, Λ, as in Eq. (4). It provides information on whether the factor risk prices significantly influence the expected returns.

The first model is the monthly CAPM with factors scaled by IP. The model marginally passes the test of HJ-distance equal zero with a p-value of 0.026. Both R_{VW} and IP are significant determinants of the correct pricing kernel, while the interaction between the two variables is not significant. Thus, the business cycle influence specified by IP is an important element missing from the CAPM. The same two factors have significant prices of risks with positive signs. Thus, a positive covariance with the market or the state of the business cycle increases the required rate of return. The fact that IP helps to explain the B/M and size effects may arise as in the framework of Jagannathan and Wang (1996) because IP could be a proxy for beta-premium sensitivity. The fact that $R_{VW} \cdot \mathrm{IP}$ is not important indicates that allowing the price of market risk to change across the business cycle is not an important determinant of the cross section of returns. Panel A of Fig. 4 reports the model's pricing errors, with its nonscaled counterpart. Most of the improvement in pricing from adding IP and $R_{VW} \cdot \mathrm{IP}$ to the CAPM occurs for low B/M portfolios, and the biggest improvement is for the smallest growth firms. As size increases, the improvement becomes smaller. However, the scaled factor model does not eliminate either the B/M or size effects. The monthly CAPM with factors scaled by IP also does not pass the supLM test at the 5% level indicating that the estimates may be unstable.

The second monthly model is the CCAPM with factors scaled by IP. Parameter estimates are reported in Panel B of Table 4. The test of HJ-distance equal zero is passed with a p-value of 0.041. The parameters associated with Δc, IP and $\Delta c \cdot \mathrm{IP}$ are all statistically significant elements of the pricing kernel. The estimates for factor risk prices indicate that both Δc and IP significantly influence the expected returns on the underlying 26 portfolios with economically sensible signs. Returns that covary positively with either consumption growth or the business cycle have higher required rates of return.

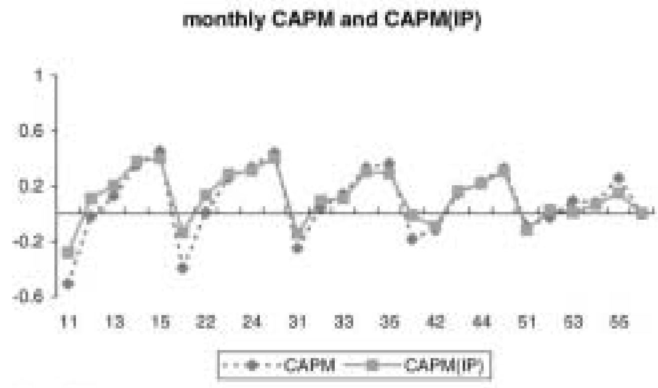
The monthly CCAPM with factors scaled by JAN also satisfies both conditions for being "interesting" with a p-value for the test of HJ-distance equal zero of 0.057. The parameter estimates are provided in Panel C of Table 4. Only the interaction between Δc and JAN is statistically significant for both the pricing kernel and prices of risk. While this result literally implies that the consumption growth rate is important only in January, an alternative interpretation is that the return characteristics of the underlying 26 portfolios are most evident in January. The pricing errors for the two scaled factor versions of the CCAPM together with the nonscaled factor benchmark are given in Panel B of Fig. 4. When the factors are scaled by IP, the improvements mostly involve a reduction of the errors for the high B/M portfolios by 0.1–0.2% per month which flattens the pricing errors relative to the nonscaled CCAPM. When the factors are scaled by JAN, both the size effect and the B/M effect are smaller and the line connecting the pricing errors is somewhat flatter.

Panel D of Table 4 reports the parameter estimates for the monthly JW model with factors scaled by IP. The *p*-value for the test of HJ-distance equal zero is 0.057. The significant determinants of the pricing kernel are R_{VW} and $R_{\text{LBR}} \cdot \text{IP}$. The same two factor risk prices along with that of $R_{\text{PREM}} \cdot \text{IP}$ significantly affect risk premiums. Panel E of Table 4 presents the parameter estimates for the monthly JW model with factors scaled by JAN. The *p*-value of the test of HJ-distance equal zero is 0.650. From the parameter estimates, both R_{LBR} and $R_{\text{LBR}} \cdot \text{JAN}$ are significant determinants of the model's pricing kernel. The parameters indicate that the factor risk price of the labor income growth rate is different in January $(-0.28 + 0.13 = -0.15)$ than outside of January (-0.28). The pricing errors of these two models together with the nonscaled JW benchmark model are presented in Panel C of Fig. 4. When the factors are scaled by IP, the pricing errors are smaller for both small firms and high B/M firms. Thus IP helps dampen both the size effect and the B/M effect. When the factors are scaled by JAN, the pricing errors are even smaller, as in the CCAPM above. However, neither of the models passes the supLM test.

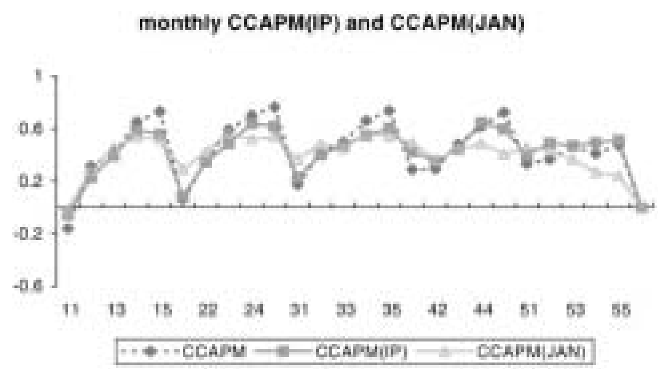
Campbell's model with nonscaled factors is reported in Panel F of Table 4. The model passes the test of HJ-distance equal zero with a *p*-value 0.347. Both the dividend yield, DIV, and the term premium, TRM, are statistically significant determinants of the pricing kernel. The second part of Panel F indicates that three variables, R_{VW}, DIV, and TRM, have statistically significant prices of risks. Neither labor income nor the relative bill rate is important in either the pricing kernel or the prices of risks. Panel D of Fig. 4

Fig. 4. Pricing errors for interesting models. The data are monthly and quarterly excess returns of the Fama-French 25 portfolios over the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12. Quarterly data are from 1953 : 01 to 1997 : 04. The portfolio numbers on the *x*-axis are numbered *ij* with *i* indexing size increasing from one to five and *j* indexing book-to-market ratio increasing from one to five. Pricing errors are defined in Eq. (27), and the numbers are in monthly (quarterly) percent.

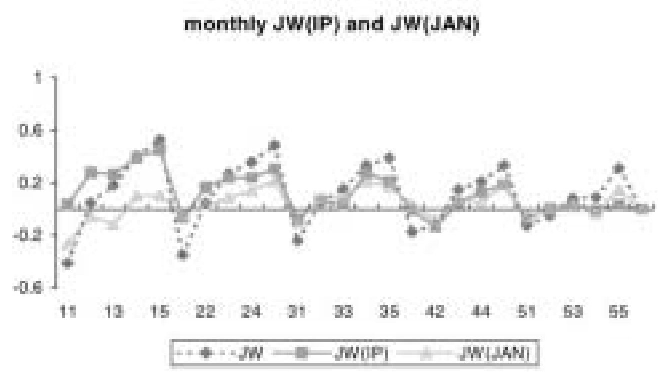
Panel A:



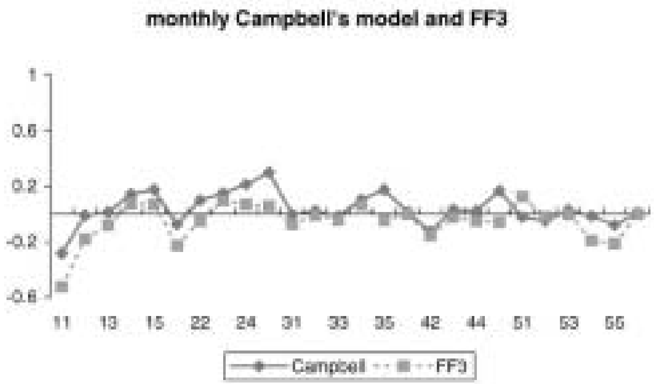
Panel B:



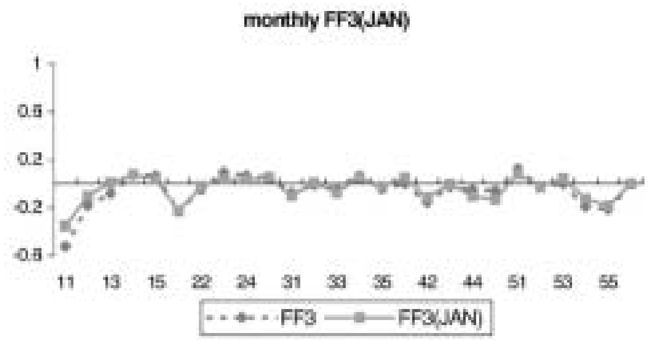
Panel C:



Panel D:



Panel E:



Panel F:

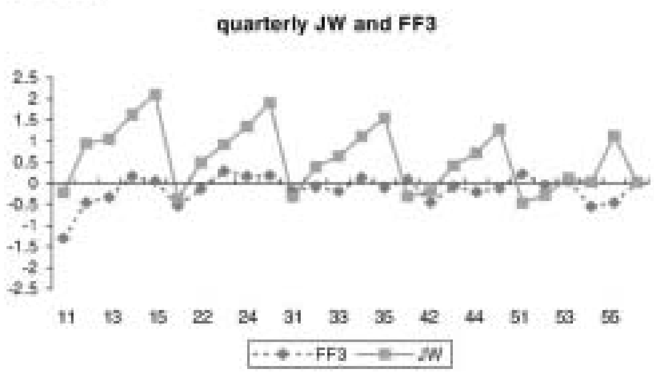
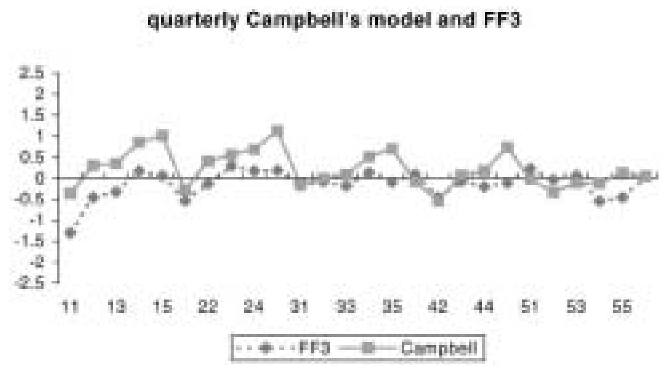
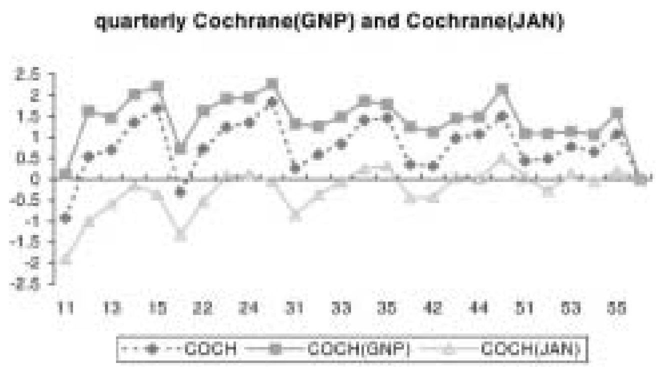


Fig. 4. (*continued*)

Panel G:



Panel H:



Panel I:

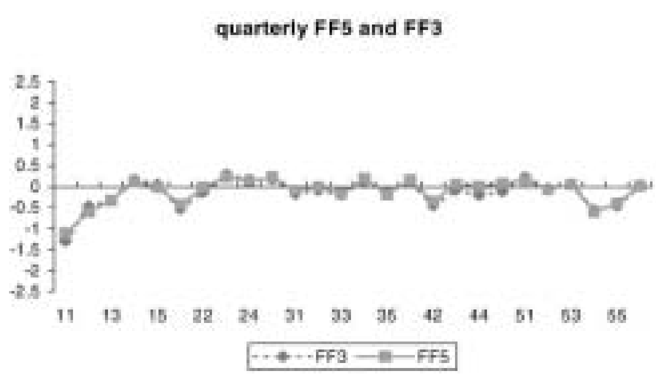


Fig. 4. (*continued*)

Table 4
Parameters estimates of interesting models

The data are returns on the Fama-French 25 portfolios in excess of the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12; quarterly data are from 1953 : 01 to 1997 : 04. The estimated parameters, $\hat{b}$, are the factor prices defined in Eq. (3). The estimated parameters, $\hat{\Lambda}$, are the beta risk prices defined in Eq. (4). The standard errors for the parameter estimates are provided in the rows labeled se.

Panel A: *Monthly CAPM with scaled factors by IP*

	Constant	R_{VW}	IP	$R_{VW}*$IP
Parameters of the pricing kernel				
$\hat{b}$	1.03	−0.04	−0.34	0.02
se	0.05	0.02	0.12	0.03
Factor risk prices				
$\hat{\Lambda}$		0.66	2.16	0.58
se		0.27	0.74	2.75

Panel B: *Monthly CCAPM with scaled factors by IP*

	Constant	Δc	IP	$\Delta c*$IP
Parameters of the pricing kernel				
$\hat{b}$	1.14	−0.75	−0.28	0.22
se.	0.10	0.36	0.11	0.12
Factor risk prices				
$\hat{\Lambda}$		0.43	1.38	−0.49
se		0.21	0.65	0.55

Panel C: *Monthly CCAPM with scaled factors by JAN*

	Constant	Δc	JAN	$\Delta c*$JAN
Parameters of the pricing kernel				
$\hat{b}$	1.05	−0.12	0.58	−3.93
se	0.06	0.37	0.90	1.62
Factor risk prices				
$\hat{\Lambda}$		0.26	0.02	0.20
se		0.22	0.06	0.08

Panel D: *Monthly JW's model with scaled factors by IP*

	Constant	R_{VW}	R_{PREM}	R_{LBR}	IP	$R_{VW}*$IP	$R_{PREM}*$IP	$R_{LBR}*$IP
Parameters of the pricing kernel								
$\hat{b}$	1.38	−0.04	−0.66	0.68	0.38	0.00	−0.40	−0.40
se	0.68	0.02	0.64	0.71	0.38	0.03	0.31	0.22
Factor risk prices								
$\hat{\Lambda}$		0.65	0.05	−0.05	1.01	0.80	1.72	1.09
se		0.28	0.12	0.13	0.98	2.68	1.02	0.41

Table 4 (*continued*)

Panel E: *Monthly JW's model with scaled factors by JAN*

	Constant	R_{VW}	R_{PREM}	R_{LBR}	JAN	$R_{VW}*$JAN	$R_{PREM}*$JAN	$R_{LBR}*$JAN
Parameters of the pricing kernel								
$\hat{b}$	−0.68	0.02	0.53	2.54	4.33	−0.45	0.35	−8.36
se	0.90	0.05	0.78	1.13	3.34	0.40	3.65	3.26
Factor risk prices								
$\hat{\Lambda}$		0.59	−0.14	−0.28	0.07	0.70	0.07	0.13
se		0.34	0.15	0.18	0.05	0.62	0.07	0.06

Panel F: *Monthly Campbell's model with nonscaled factors*

	Constant	R_{VW}	LBR	DIV	RTB	TRM
Parameters of the pricing kernel						
$\hat{b}$	−1.07	0.01	0.10	0.67	0.90	−0.72
se	1.30	0.03	0.41	0.34	4.33	0.28
Factor risk prices						
$\hat{\Lambda}$		0.66	0.02	−0.69	−0.05	1.11
se		0.31	0.27	0.33	0.04	0.35

Panel G: *Monthly FF3 with nonscaled factors*

	Constant	R_{VW}	SMB	HML
Parameters of the pricing kernel				
$\hat{b}$	1.07	−0.05	−0.01	−0.10
se	0.02	0.01	0.02	0.02
Factor risk prices				
$\hat{\Lambda}$		0.65	0.14	0.39
se		0.21	0.12	0.10

Panel H: *Monthly FF3 with scaled factors by JAN*

	Constant	R_{VW}	SMB	HML	JAN	$R_{VW}*$JAN	SMB$*$JAN	HML$*$JAN
Parameters of the pricing kernel								
$\hat{b}$	1.07	−0.08	0.12	−0.06	1.38	0.21	−0.98	0.15
se	0.05	0.03	0.06	0.05	1.22	0.26	0.43	0.42
Factor risk prices								
$\hat{\Lambda}$		0.63	0.16	0.39	0.01	−0.07	0.74	0.19
se		0.27	0.21	0.16	0.06	0.47	0.32	0.23

Panel I: *Quarterly JW's model with nonscaled factors*

	Constant	R_{VW}	R_{PREM}	R_{LBR}
Parameters of the pricing kernel				
$\hat{b}$	−0.35	0.00	−0.20	1.01
se	0.85	0.02	0.64	0.48
Factor risk prices				
$\hat{\Lambda}$		1.29	−0.02	−0.74
se		0.84	0.12	0.33

Table 4 (*continued*)

Panel J: *Quarterly Campbell's model with nonscaled factors*

	Constant	R_{VW}	LBR	DIV	RTB	TRM
Parameters of the pricing kernel						
$\hat{b}$	0.22	0.00	0.10	0.28	−0.20	−0.56
se	1.00	0.02	0.16	0.27	2.64	0.22
Factor risk prices						
$\hat{\Lambda}$		1.52	−0.13	−0.28	−0.03	0.85
se		0.79	0.37	0.24	0.02	0.34

Panel K: *Quarterly Cochrane's model with scaled factors by lag GNP*

	Constant	NRINV	RINV	GNP	NRINV∗GNP	RINV∗GNP
Parameters of the pricing kernel						
$\hat{b}$	0.92	−0.01	−0.16	0.12	−0.04	−0.09
se	0.27	0.16	0.07	0.22	0.07	0.04
Factor risk prices						
$\hat{\Lambda}$		0.33	1.76	0.03	0.86	5.33
se		0.85	1.31	0.58	1.21	3.24

Panel L: *Quarterly Cochrane's model with scaled factors by JAN*

	Constant	NRINV	RINV	JAN	NRINV∗JAN	RINV∗JAN
Parameters of the pricing kernel						
$\hat{b}$	1.41	−0.24	0.09	−1.44	0.90	−0.19
se	0.21	0.17	0.07	0.53	0.37	0.15
Factor risk prices						
$\hat{\Lambda}$		−0.63	−1.38	0.15	−1.25	−0.03
se		0.75	1.44	0.08	0.59	0.61

Panel M: *Quarterly FF5 with nonscaled factors*

	Constant	R_{VW}	SMB	HML	TERM	R_{PREM}
Parameters of the pricing kernel						
$\hat{b}$	1.23	−0.05	0.00	−0.06	−0.21	1.25
se	0.52	0.02	0.02	0.02	0.11	0.78
Factor risk prices						
$\hat{\Lambda}$		1.51	0.58	1.12	0.23	−0.06
se		0.79	0.42	0.41	0.51	0.10

reports the model's pricing errors along with the errors from the FF3 model as a comparison. No size effect is apparent and Campbell's model prices the small growth firms better than the FF3 model. While a B/M effect is present in the pricing errors of Campbell's model, its magnitude is not large. Overall, the pricing errors for Campbell's model are not bigger than those of the FF3 model, while the latter model is constructed to price the size effect and B/M effect. However, Campbell's model fails the supLM test. Thus, the parameter estimates may be unstable and should be used cautiously.

The last monthly models we report are FF3 with nonscaled factors and FF3 with factors scaled by JAN. FF3 is reported because it is so widely used, and we want to examine how it prices the size and B/M effects, which it is constructed to do. It does not pass the test of HJ-distance equal zero. Parameter estimates for FF3 are presented in Panel G of Table 4. It is somewhat surprising that only R_{VW} and HML are important for the pricing kernel, and they are also significantly priced risk factors. Panel E of Fig. 4 provides the pricing errors for FF3. The problem portfolios are the lowest B/M with smallest and second smallest sizes, which are overpriced by the model. Thus, the factor SMB cannot adequately capture the size effect in the portfolios, and SMB is not significantly priced in the unconditional version when risk prices are held constant.

The monthly FF3 with factors scaled by JAN is reported in Panel H of Table 4. It passes the test of HJ-distance equal zero with a p-value of 0.101. From the parameter estimates, R_{VW}, SMB and SMB $\cdot$ JAN are important factors for the pricing kernel. For the prices of risks, R_{VW}, HML and SMB $\cdot$ JAN are significant. This is consistent with the view that the size effect is primarily a January effect as the prices of risks for R_{VW} and HML are essentially the same across the models without and with scaling by the January dummy. As mentioned in the previous section, if the B/M effect occurred mainly in January, and HML explained the B/M effect, HML would not be priced outside January. Thus, the results tell us either there is still a significant B/M effect outside of January or there are some other risks which can be priced by HML. We also examine the pricing errors to see whether scaling by JAN really improves on the performance of the FF3 model in an interesting way. In the Panel E of Fig. 4, we find that scaling the FF3 factors with JAN actually reduces the pricing errors by 0.2% for the smallest growth stocks. Since the FF3 model already captures the B/M effect reasonably well, JAN does not improve this dimension. Both models pass the supLM test.

The first quarterly model is the JW model. It marginally passes the test of HJ-distance equal zero with a p-value 0.037. The parameter estimates are presented in Panel I of Table 4. Only R_{LBR} is statistically significant in the pricing kernel. For the prices of factor risks, R_{LBR} is also significant with a positive sign. In addition, the price of market risk is marginally significant, but R_{PREM} is not priced in contrast to Jagannathan and Wang (1996). The pricing errors of the JW model are reported in Panel F of Fig. 4 together with the quarterly FF3 model with nonscaled factors as a benchmark. Both the size effect and the B/M effect are evident in the JW pricing errors, which range from 0.5% to 2% per quarter. These pricing errors are quite large compared to those of the FF3 model. Thus, the quarterly JW model passes the HJ-distance test not because it has small pricing errors but because it has larger standard errors. Hence, our quarterly version of the JW model with nonscaled factors is not an economically interesting model. It also fails the supLM test indicating that the parameter estimates may be unstable.

The second quarterly model is Campbell's model with nonscaled factors. The test of HJ-distance equal zero has a *p*-value 0.016. Panel J of Table 4 provides the parameter estimates, and as in the monthly models, the term premium is important in the pricing kernel. Both market risk and term premium risk are priced factors for the risk premiums. The pricing errors are reported in Panel G of Fig. 4 together with the benchmark FF3. The pattern of the errors is very similar to the monthly errors in Panel D. Campbell's model improves on the smallest growth portfolio, but it has an evident B/M effect. It also fails the supLM test.

The third quarterly model is Cochrane's model with factors scaled by the cyclical element in lag GNP. The parameter estimates are given in Panel K of Table 4. For the pricing kernel, both RINV and RINV · GNP are important, and both have marginally significant prices of risks. This is consistent with Cochrane (1996) who demonstrates the importance of residential investment. The HJ-distance measure drops from 0.626 for Cochrane's nonscaled factor model to 0.559 for its scaled factor model. In all of the models discussed above, the scaled-factor models perform better than nonscaled models in the sense of HJ-distance, and we confirm that the scaling factors are economically interesting by looking at the pricing errors and parameter estimates. However, for Cochrane's model, the improvement in HJ-distance does not actually come from the improvements on pricing errors. This can be seen in Panel H of Fig. 4. The pricing errors of the nonscaled model show a distinct pattern of size and B/M effects. The scaled factor model shifts most of the pricing error upward by 0.5–1%. There is improvement only for the first portfolio. The smaller HJ-distance for the scaled factor model arises because the additional free parameters make it easier for the scaled-factor model to solve the minimization problem with the particular weighting matrix. This is significant statistically, but it is not interesting economically.

Panel L of Table 4 reports the quarterly Cochrane model with factors scaled by JAN. Both JAN and NRINV · JAN are important for the pricing kernel, and the same two factors also have significant prices of risks. By looking at Panel H of Fig. 4, we find after controlling for the January effect, the pricing errors are shifted downward by 1–1.5%, which is a big improvement for value firms. The B/M effect is mitigated but still present. Thus we conclude that the improvement in HJ-distance arises from an improvement of pricing errors. Both Cochrane's scaled factor models are stable, and they both pass the supLM test.

The quarterly FF5 model with nonscaled factors is provided in Panel M of Table 4. It passes the test of HJ-distance equal zero with *p*-value 0.018. From the parameter estimates, we find that R_{VW} and HML are determinants of the pricing kernel, as in the FF3 model, but the two macro factors, TERM and R_{PREM} are also significant determinants of the pricing kernel. The two macro factors do not have significant prices of risks. The pricing errors from FF5 in Panel I of Fig. 4 are almost the same as those in FF3. There are only small improvements on the smallest growth portfolios. Unfortunately, the two additional macro factors bring instability into the model as it fails the supLM test.

There is one last issue to note. All of the models do well in pricing the gross return of the T-bill. This implies that although the minimization problem does not put a particularly large weight on the T-bill return, it does not ignore it either. Others, such as Lettau and Ludvigson (2001b) and Jagannathan and Wang (1996), only include stock portfolios and have big estimates for the zero-beta rate. We estimate the zero-beta rate for each model. For monthly models, the rate is around 0.4% per month; for quarterly models, it is around 1.8% per quarter. We believe these estimates are more reasonable.

4.4. Values of the multipliers

We noted above that the solution for the HJ-distance from the Null model provides the least volatile element of the set of true stochastic discount factors, M. From Eq. (9) we know that $\tilde{m} = y - \tilde{\lambda}' R$, and Eq. (10) provides the estimated values of the Lagrange multipliers. The standard errors of the Lagrange multipliers are found from Eq. (24). These values for the Null model are presented in Table 5 for the monthly and quarterly data.

The Lagrange multipliers can be interpreted as portfolio weights on the basic assets. They are the product of the HJ-distance weighting matrix and the vector of average pricing errors from the model. As both the weights and the errors differ across assets and because there is correlation across the elements of the multipliers, the interpretation of the individual significance of the multipliers is best done with caution. Nevertheless, for monthly data, we find that portfolios 11,14,42, and 53 as well as the risk free return have statistically significant multipliers when the individual coefficients are evaluated at the 5% critical level. For quarterly data, these same portfolios plus portfolios 41 and 54 are also important. The importance of these portfolios is consistent with the observation that within each size quintile, there is at least one large spread position in which one of the Lagrange multipliers is a large

Table 5
λ for monthly and quarterly null models

The data are returns on the Fama-French 25 portfolios in excess of the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12; quarterly data are from 1953 : 01 to 1997 : 04. Portfolios are numbered *ij* with *i* indexing size increasing from 1 to 5 and *j* indexing book-to-market ratio increasing from 1 to 5. The Lagrangian Multipliers, λ's, are defined in Eq. (10) and their standard errors, se(λ), are defined in Eq. (26). An asterisk indicates the parameter is significant at the 5% level.

	Monthly		Quarterly	
Portfolio	λ	se(λ)	λ	se(λ)
11	−6.35*	1.73	−5.42*	1.72
12	−3.83	2.41	−3.60	2.32
13	−1.75	3.27	−3.32	3.52
14	8.76*	4.24	10.02*	4.69
15	3.72	3.71	−2.45	4.24
21	−3.66	2.65	−3.93	2.69
22	−0.09	3.15	5.02	3.32
23	6.94	3.73	5.59	3.48
24	3.40	3.75	5.28	3.83
25	2.56	3.27	4.56	3.43
31	−2.75	3.17	−3.53	3.22
32	0.02	3.67	0.17	4.05
33	−3.72	3.87	−7.36	4.36
34	5.85	3.82	4.79	4.35
35	−0.29	2.71	1.92	2.58
41	6.92	3.62	9.90*	4.03
42	−10.59*	3.95	−11.97*	4.18
43	0.09	3.66	0.91	3.98
44	−0.67	3.10	−4.63	3.64
45	0.36	2.33	2.33	2.57
51	1.78	2.43	0.28	2.39
52	−0.25	3.11	1.21	3.15
53	5.65*	2.70	5.48*	2.77
54	−4.22	2.64	−6.21*	2.85
55	−0.30	1.67	−0.16	1.72
R_f	−0.18*	0.02	−0.42*	0.06

positive number and another one close by is a large negative number. For the small firms, the portfolio positions indicate being long high B/M firms and short low B/M firms. Summing within a size quintile reveals that one would be primarily long the second and short the fourth size quintiles. Because the spread positions are probably associated with a single source of risk, it appears that there are essentially four sources of significant equity risk in these 25 portfolios.

4.5. Combining the factors of two models

An alternative way to compare models is to include the factors of several models simultaneously into the model of the pricing kernel and perform an exclusion test asking whether the second set of factors is necessary in the presence of the first. This section performs a limited comparison because the large dimensionality of the factors and scaled factor makes such a comparison impossible.

In the analysis above, both the Campbell model and the Fama-French three-factor model are reasonably successful. By including the two additional FF3 factors, SMB and HML, in the pricing kernel of the Campbell model, one can test whether they are significant additional determinants of the pricing kernel. The results of this analysis are presented in Panel A of Table 6. Notice that none of the individual coefficients is significant at the 0.05 level of significance, in strong contrast to the results of the individual models. This is an indication of multicollinearity. Correlation across the factors also makes the exclusion tests inconclusive. The *p*-value of the Wald test that the parameters associated with SMB and HML are zero is 0.135 indicating that these factors are unnecessary once the Campbell factors are present, but the comparable test that the FF3 model does not need the four additional factors of Campbell's model has a *p*-value of 0.215. Thus, since the factors of the respective models are significant when included individually, we can conclude that the same basic information is captured in different ways by the two models.

To avoid problems with multicollinearity, Campbell (1996) orthogonalizes the factors and scales them to have the same variance as the market return. The first factor is the market return, the second is the part of labor income that is not explained by the market return, the third is the part of the dividend yield that is not explained by the market return and labor income, and so on. When we place the two Fama-French factors after the five Campbell factors, we ask whether the parts of SMB and HML that cannot be explained by the Campbell factors are significant determinants of the pricing kernel. The results are presented in Panel B of Table 6.

The coefficients on R_{VW}, DIV, TRM, and HML are all more than 1.5 times their standard errors. In particular, even though HML is placed last in the ordering of variables, its *p*-value remains 0.069. Thus, HML appears to add some independent information to the pricing kernel over and above that provided by the Campbell factors.

Panels C and D of Table 6 report the results of a hybrid model that uses these four elements with orthogonalized factors. The hybrid model has the smallest HJ-distance, 0.285, of any of the estimated models, and the tests indicate no evidence against the model, except for the stability test which again indicates potential problems with the model.

Table 6
Combining factors of Campbell's model and the Fama-French three-factor model

The data are returns on the Fama-French 25 portfolios in excess of the T-bill rate and the return on the T-bill. Monthly data are from 1952 : 01 to 1997 : 12; quarterly data are from 1953 : 01 to 1997 : 04. The factors are collected from Campbell's model and FF3. We use Cholesky decomposition to orthogonalize factors in Panels B and C. The parameter estimates, $\hat{b}$, are factor prices defined in Eq. (3). The standard errors for $\hat{b}$ are provided in the row of se. The *p*-value for the test $\hat{b} = 0$ is $p(\hat{b} = 0)$. HJ-dist(δ) is Hansen-Jagannathan distance. *p*-value for the test $\delta = 0$ calculated under the null $\delta = 0$ is $p(\delta = 0)$. Max. Error is the maximum annual pricing error for a portfolio with annual standard error of 20% under the assumption $E(m) = E(y)$. The standard error for HJ-distance under the alternative hypothesis $\delta \neq 0$ is se(δ). The *p*-value of the optimal GMM test is $p(J)$. The value of supLM statistics is supLM. An asterisk indicates the model fails the supLM test at the 5% significance level. No. of para is the number of parameters.

Factors	Constant	R_{VW}	LBR	DIV	RTB	TRM	SMB	HML
Panel A: Factors prices for the combined model (factors are not orthogonalized)								
$\hat{b}$	−0.31	−0.02	−0.11	0.43	0.70	−0.38	−0.02	−0.06
se	1.03	0.02	0.33	0.27	3.33	0.26	0.02	0.03
$p(\hat{b} = 0)$	0.76	0.35	0.74	0.11	0.83	0.14	0.37	0.07
Panel B: Factors prices for the combined model (factors are orthogonalized)								
$\hat{b}$	−0.31	−0.05	−0.03	0.11	0.07	−0.10	−0.01	−0.03
se	1.03	0.01	0.07	0.06	0.07	0.06	0.01	0.02
$p(\hat{b} = 0)$	0.76	0.00	0.61	0.09	0.27	0.12	0.46	0.07

Factors	Constant	R_{VW}	DIV	TRM	HML
Panel C: Factors prices for the hybrid model (factors are orthogonalized)					
$\hat{b}$	0.02	−0.05	0.09	−0.14	−0.03
se	0.95	0.01	0.06	0.06	0.02
$p(\hat{b} = 0)$	0.99	0.00	0.12	0.02	0.11

HJ-dist(δ)	$p(\delta = 0)$	Max. error	se(δ)	$p(J)$	SupLM	No. of para
Panel D: Summary statistics for the hybrid model						
0.285	0.235	5.7%	0.058	0.144	192.736	5

The intuition of the Campbell model is that any variable that predicts the market return in a multivariate setting is a potential factor that affects the cross-section of asset prices. To determine whether HML arises as a risk factor within this restricted context we estimated a vector autoregression of the four factors. The results indicate that HML is not an important determinant of the other three variables because the smallest p-value associated with the coefficients on HML in any of the three equations was 0.336. The HML equation also indicated that none of the other three variables is a significant determinant of HML, although there is evidence of own serial correlation. Thus, if HML is a risk factor, its importance must be traced to the more general economic state variables of Merton (1973) rather than the restrictions arising in Campbell's model. Some support for this position is provided by Liew and Vassalou (2000) and Vassalou (2000) who argue that SMB and HML are risk factors that arise because of their ability to predict future GDP.

4.6. Robustness

In all of the above results, we obtain parameter estimates and conduct tests using nonscaled returns. To examine whether these models are robust, we change the underlying assets from nonscaled returns to scaled returns, and we investigate whether the parameter estimates obtained from nonscaled return models (the first stage estimates) can price the scaled returns. We scale returns with the term premium, the difference in yields between a 30-year government bond and a one-year government bond. If a model is able to price the basic assets (nonscaled returns), and it is specified correctly, it should be able to price the scaled returns, which can be thought of as a set of managed portfolios in which the manager invests different amount depending on the realization of the term premium.

Table 7 provides the information on these experiments. We use the estimates obtained from the first stage by optimal GMM, to calculate the test of the HJ-distance equal zero for the scaled returns and the J-statistic for optimal GMM for the new orthogonality conditions, as in Eq. (30). These p-values are denoted p_1 and p_2. We also use the first-stage estimates of HJ-distance to calculate second-stage HJ-distance tests, and the p-value is denoted p_3. None of the monthly models successfully prices the new assets.

5. Conclusion

This paper evaluates a number of asset pricing models proposed in light of the anomalies uncovered in testing the CAPM. The models are compared on a common set of returns: 25 size and book-to-market portfolios constructed as in Fama and French (1993) for a sample period from 1952 to 1997. Average excess returns across these portfolios are as low as 0.36% per month and as high as 1.13% per month. Within a size quintile, higher book-to-market portfolios have higher average returns. Within all but the lowest

Table 7
Robustness test for nonscaled returns models

The tests are based on returns on the Fama-French 25 portfolios in excess of the T-bill rate and the return on the T-bill, conditioned on the term premium, the difference in yields between a 30-year government bond and a one-year bond. Monthly data are from 1952 : 01 to 1997 : 12; quarterly data are from 1953 : 01 to 1997 : 04. The p-values are: $p1$, test of HJ-distance $= 0$ using parameter estimates from optimal GMM for corresponding nonscaled return models; $p2$, test of optimal GMM over-identification using parameter estimates from optimal GMM for corresponding nonscaled return models; $p3$, test of HJ-distance $= 0$ using parameter estimates from minimizing HJ-distance for corresponding nonscaled return models.

	NULL	CAPM	CCAPM	JW	CAMP	FF3	FF5
Panel A: Monthly scaled returns by TERM with nonscaled factors							
$p1$	0	0	0	0	0	0	0
$p2$	0	0	0	0	0	0.001	0.003
$p3$	0	0	0	0	0	0	0
Panel B: Monthly scaled returns by TERM with scaled factors by IP							
$p1$	0	0	0	0.002	0	0	0
$p2$	0	0.004	0	0.004	0	0.017	0
$p3$	0	0	0	0	0	0	0
Panel C: Monthly scaled returns by TERM with scaled factors by JAN							
$p1$	0	0	0.001	0	0	0	0
$p2$	0	0.001	0.036	0.002	0	0.007	0.086
$p3$	0	0	0.075	0.004	0	0.001	0.001

	NULL	CAPM	CCAPM	JW	CAMP	COCH	FF3	FF5
Panel D: Quarterly scaled returns by TERM with nonscaled factors								
$p1$	0	0	0.014	0.002	0	0	0.002	0.003
$p2$	0.003	0.005	0.012	0.006	0	0	0.040	0.049
$p3$	0	0	0.006	0.001	0	0	0.001	0.002
Panel E: Quarterly scaled returns by TERM with scaled factors by lag GNP								
$p1$	0	0.002	0.010	0.001	0	0	0.039	0.018
$p2$	0.001	0.017	0.015	0.008	0	0.001	0.361	0.495
$p3$	0	0.002	0.004	0	0	0	0.069	0.016
Panel F: Quarterly scaled returns by TERM with scaled factors by CAY								
$p1$	0	0.001	0	0.005	0	0	0.001	0.001
$p2$	0.001	0.006	0.002	0.012	0	0.001	0.053	0.110
$p3$	0	0	0	0.003	0	0	0.002	0.006
Panel G: Quarterly scaled returns by TERM with scaled factors by JAN								
$p1$	0	0.002	0.001	0.020	0	0.064	0	0.006
$p2$	0.015	0.020	0.011	0.103	0.017	0.236	0.032	0.216
$p3$	0.001	0.002	0	0.016	0	0.039	0.001	0.004

book-to-market quintiles, average returns are generally decreasing in size. The unconditional CAPM cannot explain these returns.

We consider only linearized versions of the models, and we evaluate the models with both nonscaled factors and scaled factors, where the scaling reflects either business-cycle movements or a January dummy. The models are compared using the methodology of Hansen and Jagannathan (1997) who recognize that the estimated distance between a model's pricing kernel and the true pricing kernel also is an estimate of the maximal mispricing of a portfolio of the assets. We also evaluate the models using the optimal GMM test of Hansen (1982). In general, we find little disagreement between the two tests. Finally, we evaluate the temporal stability of the parameters using the supLM test of Andrews (1993).

For monthly models with nonscaled factors, Campbell's (1996) model is the only model that passes the test of HJ-distance equals zero, and its estimated HJ-distance is also smaller than that of the Fama-French (1993) three-factor model. Only three of the five factors in the model appear to be important: the return on the market portfolio, the dividend yield, and the term premium. The HML factor of the Fama-French model also has independent information over and above that provided by these three factors. Unfortunately, the Campbell model fails to pass the stability test. While the simulation study of Ahn and Gadarowski (1999) provides some support that the small-sample distributions of the HJ-distance test are reliable for our sample size, no comparable study of the small-sample distributions of the stability test has been conducted. Thus, additional study of the Campbell model is desirable. In particular, we evaluate only the linearized version of the model.

Scaling the risk factors of the models with the cyclical element in industrial production as measured by the Hodrick–Prescott (1997) filter improves the performance of several of the models. The CAPM, CCAPM, and Jagannathan and Wang (1996) models all have significant coefficients on the scaled factors. There is also evidence that pricing in January is significantly different than pricing outside of January. For example, when the three factors of the Fama-French (1993) model are entered without scaling, only the market return and the HML portfolio are significant risk factors. When the factors are also scaled with a January dummy, the market return and the HML portfolio retain their significance and the SMB portfolio is significant in January. This latter model also passes the stability test.

With quarterly data, none of the models with nonscaled factors passes the test of HJ-distance equal to zero. Nevertheless, the simulation results of Ahn and Gadarowski (1999) suggest that these results should be interpreted with care as the sizes of the tests appear to deteriorate in this sample size. Neither scaling with the cyclical component of GNP, as measured by the Hodrick–Prescott (1997) filter, nor scaling with the consumption-wealth series of Lettau and Ludvigson (2001a) has much of an influence on the results.

Additionally, none of the models, either monthly or quarterly appears to be robust in the following sense. When we estimate the parameters of the models using the basic returns and ask the models to price the set of assets constructed by scaling returns with the term premium, all of the models fail.

There are several directions in which this study could be extended. First, we construct our estimates as if there are no transactions costs or short-sale constraints in asset markets. Hanna and Ready (1999) find that transaction costs reduce but do not eliminate the CAPM anomalies. Luttmer (1996) notes that small transaction costs and short-sale constraints can have large implications for the variability of implied stochastic discount factors. Recall that Fig. 1 indicates that the HJ-distance methodology requires the models to price large short-sale positions. Future research should be directed to determine how transaction costs and short-sale constraints affect the estimates of HJ-distance. Liquidity and market impact of trading individual assets may also be important. Brennan et al. (1998) find that average returns on individual equities are affected by trading volume, which is consistent with differences in liquidity premiums across assets. Understanding how liquidity is priced and the role it plays in portfolio returns is an open issue. The presence of these market frictions implies that it may be difficult if not impossible to realize the returns that certain trading strategies imply. It is only truly available returns that require adjustment for risk.

Appendix A. Distribution of HJ-distance

The distribution of δ is not standard under the assumption that the true δ equals zero. Jagannathan and Wang (1996) demonstrate that the distribution of $T\delta^2$ involves a weighted sum of $n-k\chi^2(1)$ statistics, where n is the number of assets and k is the number of estimated parameters. The weights are the $n-k$ nonzero eigenvalues of

$$A = S_T^{1/2} W_T^{1/2\prime}[I_n - W_T^{1/2} D_T (D_T' W_T D_T)^{-1} D_T' W_T^{1/2\prime}] W_T^{1/2} S_T^{1/2\prime}, \qquad \text{(A.1)}$$

where $S_T^{1/2}$ and $W_T^{1/2}$ are the upper-triangular matrices from the Cholesky decompositions of S_T and W_T, and I_n is the n-dimensional identity matrix. It can be demonstrated that A has exactly $n-k$ nonzero eigenvalues, which are positive and are denoted by $\theta_1, \ldots, \theta_{n-k}$. Then, the asymptotic sampling distribution of the HJ-distance is

$$T\delta^2 \xrightarrow{\mathrm{d}} \sum_{j=1}^{n-k} \theta_j v_j \quad \text{as } T \to \infty, \qquad \text{(A.2)}$$

where $v_1, \ldots, v_{n-k}$ are independent $\chi^2(1)$ random variables. We simulate the statistics 10,000 times to determine the p-value for the estimated HJ-distance.

Appendix B. Robustness check with scaled returns

We first calculate parameter estimates from optimal GMM using the 26 returns as

$$\hat{b} = \arg\min g_T(R, b)' W^* g_T(R, b). \tag{B.1}$$

Then, under the null that $\hat{b}$ is the true parameter, the set of scaled returns Rx should be correctly priced with $\hat{b}$. We calculate the new J statistics as

$$J = g_T(Rx, \hat{b})' var[g_T(Rx, \hat{b})]^{-1} g_T(Rx, \hat{b}), \tag{B.2}$$

where

$$g_T(Rx, \hat{b}) = \frac{1}{T} \sum_{t=1}^{T-1} [(R_{t+1} x_t)(\hat{b}' F_{t+1}) - p x_t]. \tag{B.3}$$

The J-statistic is distributed as a $\chi^2(n)$ under the null. The degrees of freedom are n because we have n orthogonality conditions, and we do not estimate any additional parameters. The same argument applies to HJ-distance. With the new orthogonality conditions for scaled returns, we need to calculate the new δ and the distribution of $T\delta^2$. Since the first stage estimates by optimal GMM are not very different from those obtained from HJ-distance estimation, we choose to use the estimates from optimal GMM to calculate new HJ-distances for the new scaled assets.

References

Ahn, S., Gadarowski, C., 1999. Small sample properties of the model specification test based on the Hansen-Jagannathan distance. Unpublished working paper, Arizona State University.

Andrews, D., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.

Breeden, D., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.

Brennan, M., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373.

Campbell, J., 1993. Intertemporal asset pricing without consumption data. American Economic Review 83, 487–512.

Campbell, J., 1996. Understanding risk and return. Journal of Political Economy 104, 298–345.

Campbell, J., Cochrane, J., 2000. Explaining the poor performance of consumption-based asset pricing models. Journal of Finance 55, 2863–2878.

Cochrane, J., 1996. A cross-sectional test of an investment-based asset pricing model. Journal of Political Economy 104, 572–621.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. Journal of Finance 52, 1–33.

Daniel, K., Torous, W., 1995. Common stock returns and the business cycle. Unpublished working paper, Northwestern University.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Fama, E., French, K., 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131–155.
Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55–84.
Ferson, W., Harvey, C., 1999. Conditioning variables and the cross section of stock returns. Journal of Finance 54, 1325–1360.
Ghysels, E., 1998. On stable factor structure in the pricing of risk: Do time-varying betas help or hurt? Journal of Finance 53, 549–573.
Glasserman, P., Jin, Y., 1998. Comparing stochastic discount factors through their implied measures. Unpublished working paper, Columbia University.
Hanna, J., Ready, M., 1999. Profitable predictability in the cross-section of stock returns. Unpublished working paper, University of Wisconsin.
Hansen, L., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.
Hansen, L., Singleton, K., 1982. Generalized instrumental variable estimation of nonlinear rational expectation models. Econometrica 50, 1269–1286.
Hansen, L., Jagannathan, R., 1997. Assessing specification errors in stochastic discount factor models. Journal of Finance 52, 557–590.
Hodrick, R., Prescott, E., 1997. Postwar U.S. business cycles: an empirical investigation. Journal of Money, Credit and Banking 29, 1–16.
Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51, 3–53.
Lettau, M., Ludvigson, S., 2001a. Consumption, aggregate wealth and expected stock returns. Journal of Finance 56, 815–849.
Lettau, M., Ludvigson, S., 2001b. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. Journal of Financial Economics, forthcoming.
Liew, J., Vassalou, M., 2000. Can book-to-market, size, and momentum be risk factors that predict economic growth? Journal of Financial Economics 57, 221–245.
Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.
Loughran, T., 1997. Book-to-market across firm size, exchange, and seasonality: is there an effect? Journal of Financial and Quantitative Analysis 32, 249–268.
Luttmer, E., 1996. Asset pricing in economies with frictions. Econometrica 64, 1439–1467.
Merton, R., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.
Roll, R., 1977. A critique of the asset pricing theory's test: part I: on past and potential testability of the theory. Journal of Financial Economics 4, 129–176.
Sharpe, W., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442.
Vassalou, M., 2000. News related to future GDP growth as risk factors for equity returns. Unpublished working paper, Columbia University.